Research Note 84-119

HUMAN RESOURCES TEST AND EVALUATION SYSTEM (HRTES) Comprehensive Handbook

Jonathan D. Kaplan, William H. Crooks,
Mark S. Sanders, and Rina Dechter

Perceptronics, Inc.

Contracting Officer's Representatives
Irving N. Alderman and Charles O. Nystrom

Battlefield Information Systems Technical Area
Franklin L. Moses, Chief

SYSTEMS RESEARCH LABORATORY
Jerrold M. Levine, Director

U. S. Army
Research Institute for the Behavioral and Social Sciences

August 1984

Approved for public release; distribution unlimited.

U. S. ARMY RESEARCH INSTITUTE

FOR THE BEHAVIORAL AND SOCIAL SCIENCES

A Field Operating Agency under the Jurisdiction of the

Deputy Chief of Staff for Personnel

L. NEALE COSBY
Colonel, IN
Commander

EDGAR M. JOHNSON
Technical Director

1. INTRODUCTION

I. Overview.

This handbook is designed to assist you in evaluating the performance of the operators and maintainers in a system. The Human Resources Test and Evaluation System (HRTES) is a set of procedures which aids you in applying your military judgment to (1) develop the test plans which will test the critical aspects of human performance, (2) evaluate that performance, and (3) diagnose the cause(s) of inadequate performance. These causes may be inadequate training, human factors engineering, or manpower selection. The procedures for planning an operational test and for evaluating and diagnosing the test results are described in this Handbook. The accompanying Workbook includes all of the HRTES forms and instructions that you will need during test planning, evaluation, and diagnosis. [illegible] weapon systems [illegible]

Before embarking on a description of the HRTES procedures and their relation to existing test and evaluation procedures, it is important to discuss a basic assumption in HRTES. In developing HRTES we have assumed that the primary purpose for conducting test and eveluation is to determine whether the tested system is able to satisfy the requirements for which it was developed. Since systems are developed to perform activities, satisfying the requirements implies that the system is able to perform all of those activities necessary to fulfill these requirements. Given this assumption, the HRTES procedures lead you to focus first on identifying the activities that the system must perform. Since the emphasis of HRTES is on the human components of the system, the HRTES procedures next lead you to identify those human activities which must be performed if the system as a whole is to be able to perform its overall activities. Only after the required human activities have been identified, do you then consider "what to measure." Finally, HRTES includes procedures to assist you to evaluate the test results after the field data have been collected and to diagnose probable causes of inadequate human performance.

HRTES has been designed to complement the existing OT&E guidelines described in the Force Development Operational Testing and Evaluation Methodology and Evaluation Methodology and Prcoedure Guide (AR 71-3). HRTES was designed to meet the reporting requirements that, according to AR 71-3, are a part of the OT&E cycle. These reports are: (1) the Independent Evaluation Plan (IEP), (2) the Outline Test Plan (OTP), (3) the Test Design Plan (TDP), (4) Detailed Test Plan (DTP), (5) the Test Report (TR), and (6) the Independent Evaluation Report (IER). Chapters two through six of HRTES utlimately result in a test plan for OT and, as such, provide material for the IEP, OTP, TDP, and the DTP. Chapters seven and eight of HRTES describe procedures for system evaluation and diagnosis and thus yield the material necessary for completing the TR and the IER. Moreover, HRTES is designed to be general enough to be applied to both major and nonmajor systems in all operational tests.

II. Purpose for HRTES Development.

As a result of increasing weapon complexity, increasing demands for highly trained operators, and higher costs for materiel and personnel, material testing programs are under severe pressure to assure the timely detection and evaluation of potential system problems. Undetected problems due to failures in maintainability and reliability lead to system unavailability and increased life cycle costs. In addition, the knowledge that 50 to 70 percent of all failures of major weapons and space systems are caused by human-initiated failures underscores the importance of including human resource considerations throughout the system acquisition process (Howard and Lipsett, 1976).

The contribution of system operators and maintainers to system performance is often more difficult to assess than system hardware and software components. There are a number of reasons why the human element poses great difficulty in system evaluation. The increasing complexity and sophistication of weapon systems make greater and greater behavioral demands of system

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: ARI Research Note 84-119	2. GOVT ACCESSION NO.
3. RECIPIENT'S CATALOG NUMBER	
4. TITLE (and Subtitle): HUMAN RESOURCES TEST AND EVALUATION SYSTEM (HRTES) Comprehensive Handbook	5. TYPE OF REPORT & PERIOD COVERED: Final Report period ending May 1982
	6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s): Jonathan D. Kaplan, William H. Crooks, Mark S. Sanders and Rina Dechter	8. CONTRACT OR GRANT NUMBER(s): DAHC 19-77-C-0055
9. PERFORMING ORGANIZATION NAME AND ADDRESS: Perceptronics, Inc. 6271 Variel Avenue Woodland Hills, CA 91367	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: 2Q262717A765 and 2Q263743A775
11. CONTROLLING OFFICE NAME AND ADDRESS: U.S. Army Research Institute for the Behavioral and Social Sciences, 5001 Eisenhower Avenue, Alexandria, VA 22333-5600	12. REPORT DATE: August 1984
	13. NUMBER OF PAGES: 236
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office): --	15. SECURITY CLASS. (of this report): Unclassified
	15a. DECLASSIFICATION DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

--

18. SUPPLEMENTARY NOTES

This is the comprehensive version of HUMAN RESOURCES TEST AND EVALUATION SYSTEM (HRTES). It differs in many respects from the later HRTES TEST PROCEDURES AND SUPPLEMENT. This is Volume I of a two-volume set. (over)

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Test Evaluation
Performance Taxonomy
Performance Testing
Criteria
Operational Testing
Human Factors Testing

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

This research note is the first volume of a two-volume set designed to aid in making the decisions needed in operational testing, including front-end analysis The series of structured decision aids herein aid in determining the required classes of performance, the conditions that apply to performance, the criteria for performance, the measures of performance, the value of performance outcomes, and the causes of inadequate system performance.

DD FORM 1473, 1 JAN 73 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

ARI Research Note 84-119

13. (continued)

Volume II is published seperately as ARI Research Note 84-120. See also related ARI Research Products 84-19 and 84-20.

Irving N. Alderman and Charles O. Nystrom, contracting officer's representatives

FOREWORD

There are currently two versions of the Human Resources Test and Evaluation System (HRTES) available. This is the original, comprehensive version. This version was designed to aid all three major test and evaluation processes in the most methodologically sound manner available, without compromise. These three processes are:

(1) Front-End Analysis;
(2) Test Plan Development;
(3) Evaluation of Test Results.

The second version of HRTES (HRTES Test Procedures and Supplement) was designed to aid only the latter two of these processes. This second version provides test and evaluation methods that (in general) require substantially less time and interorganizational cooperation than do those of the original HRTES.

This (the comprehensive) version of HRTES is structured around the creation of a hierarchical model of system performance, the definition of highly specific criteria of performance and the measures of that performance, and the determination of the criticality of the various elements of this hierarchical model. HRTES differs from other test and evaluation systems in actually providing: candidate elements at each of the hierarchical levels, methods for rating these elements, and the actual forms for such rating.

The use of HRTES is dependent upon two general factors: the part of the T&E cycle you are currently involved with, and the nature and extent of the aid you feel that you require. Since HRTES attempts to aid a great deal of the T&E cycle, you may wish to use only part of it. Chapter One should help you to understand HRTES' structure so that you can decide which chapters are most likely to be helpful to you. You should be aware that HRTES was written at an operational level to the greatest extent possible. That is, it describes not only what should be done, but how to do it. Because of this operational orientation, it may appear that to use HRTES you must follow all of its procedures and methods in "lock step." This is not intended. HRTES is an aid. If you follow it exactly as written, it should help you produce a highly logical portion of the T&E cycle that will result in a well-reasoned series of acquisition decisions. However, you may use any parts of HRTES in the way you think will best aid you. You may exerpt tasks or conditions or methods for aiding the development of criteria as you see fit. To the extent that you agree with the methods in HRTES and have the time to use them, they should prove useful. To the extent that you can use parts of HRTES according to another T&E method, you should feel free to do so.

HRTES is a new method. In the last analysis, to be fully useful it will have to alter or grow according to the changing realities of the acquisition cycle. To provide for this improvement, a so-called HRTESGRAM has been included on the second page of the HRTES Workbook. Your input and ideas are greatly desired. If you have any suggestions, please use this form (and the address contained therein) to communicate them.

TABLE OF CONTENTS

TABLE OF CONTENTS
(Continued)

GLOSSARY

1. AGGREGATION ACROSS CONDITION SETS:

This means the statistic for a given Human Performance Function is collected for all the condition sets of that HPF. This would be done if conditions were included only to provide a representative situation for testing, not if one wanted to know the effects of specific conditions on HPF performance. Agrregation across condition sets also reduces the number of trials required to assure performance reliability.

2. AGGREGATION WITHIN EACH CONDITION SET:

This means the statistic for a given Human Performance Function is collected separately for each condition set of that HPF. This would be done if one wanted to know the effects of specific conditions on HPF performance, not if conditions were included only to provide a representative situation for testing. Aggregation within each condition set of a given HPF will significantly increase the number of trials of that HPF which are required to assure performance reliability.

3. BRANCHES:

Branches connect nodes in one level of a Tree with appropriate, related nodes in the adjacent level(s). For example, a given System Performance Issue Node will be connected by branches to one Superordinate System Function Node and its related Subordinate HPF-Group Nodes.

4. CONDITION:

A condition is any element which will significantly effect the ability of a system (including its human components) to perform a given System Performance Issue or those Human Performance Functions required by that Issue. In general, conditions are variables of: a system's environment, a system's operational status, a system's tactics/behavior, a system's preparation, the nature of a system's enemy/target, the tactics/behavior of a system's enemy/target. In HRTES, conditions may be included in System Performance Issues to modify them. They are also used, in a highly detailed manner, to modify the performance of each Human Performance Function.

5. CONDITION CATEGORY:

In HRTES, conditions are collected in organized categories in the "Condition Category Index." Condition categories are rated for criticality, and if selected based on criticality rating, the individual conditions within that category are then rated.

6. CONDITION SET:

A condition set is a collection of conditions under which the measures of a Human Performance Function will be taken. A condition set, by definition, cannot have two conditions from the same condition category. A condition set consists of one condition from each category that was rated critical. Condition sets only apply to Human Performance Functions, but they contain all conditions from a given HPF's parent System Performance Issue. A Human Performance Function may be performed under more than one condition set.

7. CRITICAL INCIDENT:

A critical incident is an accident, or near accident, which either produced or might have produced significant damage to personnel or hardware. Such an incident must be recorded and reported immediately to appropriate test authorities.

8. CRITICALITY RATING:

To select the types of performance which must be tested, criticality ratings of competing performance types are done in HRTES. Criticality can be rated as a global scale, or it can be divided into those attributes which constitute the whole, with each attribute being rated separately. In general, HRTES follows the latter approach.

9. EVALUATION WEIGHT:

Evaluation weights are the result of manipulation of the selection weights. They are included in the nodes of the Evaluation Tree, and are used in evaluating performance following the field test.

10. EXPERTS:

Experts are individuals who have a significant knowledge of a specified functional area (e.g., system requirements, system operations, system maintenance, tactics) and/or technical area (training, Human Factors Engineering, manpower selection). They are used extensively in the HRTES procedures. By definition, the test planner/evaluator is an expert.

11. HANDBOOK:

The handbook is the volume of HRTES which contains the instructions and other information for the HRTES user which do not have to be copied/duplicated for use.

12. FOLDING BACK THE EVALUATION TREE:

This is the procedure in which values, developed in Value Functions, and evaluation weights are manipulated mathemtically to produce appropriate values for higher level nodes of the Evaluation Tree.

13. HUMAN PERFORMANCE FUNCTION (HPF):

HPF's are specific activities which humans must perform to enable the system to perform a System Performance Issue. In effect, HPF's are tasks which have been developed to be applicable to functionally related systems. There are two general types of HPF, Operational HPF's and Maintenance HPF's.

14. HUMAN PERFORMANCE FUNCTION-GROUPS (HPF-GROUPS):

An HPF-Group is a collection of those Human Performance Functions which must be performed together so that a related System Performance Issue can be performed. There are two general types of HPF-Group, Operational HFP-Groups and Maintenance HPF-Groups for a gievn System Performance Issue.

15. HUMAN RESOURCE/DIAGNOSTIC MEASURE:

These are diagnostic measures of the Human Resource Areas: Training, Human Factors Engineering, and Manpower Selection. These measures are used to determine the probable cause(s) of inadequate performance of each significant Human Performance Function.

16. INDEX OF ACCEPTABILITY:

The measures of a given Human Resource Area (Training, Human Factors Engineering, and Manpower Selection) can be collapsed to produce an Index of Acceptability for that Human Resource Area in relation to a specified Human Performance Function. This Index always is a number between zero and one hundred, with one hundred meaning that for the HPF, that Human Resource Area was dealt with in a completely acceptable manner.

17. LEVEL OF A TREE:

Each level of a Selection or Evaluation Tree consists of nodes of the same type. These trees in HRTES have six levels, one for each type of node.

18. LOA:

Letter of Agreement.

19. MENS:

Mission Element Needs Statement.

20. NODES:

Nodes are those parts of the Selection and Evaluation Trees which represent the major performance requirements of the system. There are six levels of nodes in a Tree: System Node, System Function Node(s), System Performance Issue Node(s), Human Performance Function-Group Node(s), Human Performance Function Node(s), and Statistic Node(s).

21. OBSERVERS:

Observers are those field test personnel whose job it is to collect performance and diagnostic data during the test.

22. PARTICIPANTS:

Participants are those personnel who will be operating and maintaining the systems' hardware.

23. PERFORMANCE CRITERION:

A performance criterion is the definition of one successful trial of one Human Performance Function as performed under one condition set. It is formated in terms of maximum permitted time and/or minimum permitted accuracy plus a statement of the effect of an accident or near accident on a trial. Accuracy, itself, is stated in terms of number/percentage of errors, with the occurance of each error defined.

24. PERFORMANCE UNIT:

In HRTES, one performance unit consists of those individuals required to perform one trial of a given Human Performance Function. Thus, one performance unit for detecting targets with a rifle might be one individual, whereas one performance unit for detecting targets with a medium tank might be two individuals (commander and gunner).

25. PERFORMANCE VALUE:

Performance value is a number from zero to one hundred which expresses an expert's value of the statistic/performance of a Human Performance Function. It is read from the Value Scale of "Value Function Work-Sheet" at the point where a Value Function line intersects the Scale.

26. RELATIVE WEIGHT:

Relative weight is a normalized weight which applies to each System Performance Issue Node of a given family of nodes. A relative weight is a number between zero and one. All relative weights of a specific family of SPI Nodes sum to one.

27. RELIABILITY:

Reliability is the extent to which the one performance by an individual is representative of that individual's performance repeated over times. Reliability is also the extent to which the performance of a sample of individuals would be repeated by other samples from the same population.

28. ROC:

Required Operational Characteristics.

29. ROOT:

The root of a Selection or Evaluation Tree is the node from which all other nodes are decomposed, in this case a System Node. It is normally portrayed at the top of a Tree with all other nodes branching down from it.

30. SELECTION TREE:

The Selection Tree is a structure consisting of six levels of nodes connected by branches. The levels are: System Level, System Function Level, System Performance Issue Level, Human Performance Function-Group Level, Human Performance Function Level, and Statistic Level. Each node contains a selection weight. The Selection Tree serves as both an audit trail which is filled in by the HRTES user and a basic tool for the eventual evaluation of performance.

31. SELECTION WEIGHT:

Selection weights are the result of the criticality rating of competitive elements in HRTES leading to the selection of some and the rejection of others. These weights are included in the nodes of the Selection Tree.

32. STATISTIC:

In HRTES, a statistic is the data from multiple trials of a given Human Performance Function. In HRTES, two types of statistics may be chosen, percentage of successful trials, or average time or accuracy. A statistic may be collected for each condition set of a given Human Performance Function, or across all the condition sets of a Human Performance Function.

33. STATISTIC CRITERION:

A statistic criterion is the definition of success for the performance of multiple trials of a given Human Performance Function. This definition will be based on the type of statistic used: percentage or average. A percentage criterion may be viewed as equivalent to a probability criterion. That is if a probability criterion were--"80% probability of successful performance of one HPF trial" then the equivalent percentage criterion would be--"80% of trials of the HPF must be performed successfully." A statistic criterion may be developed for each condition set of a given HPF, or across all condition sets of an HPF.

34. SYSTEM:

The word "system" refers to the military system to be tested. It includes: hardware; software, if any; operators; and maintenance personnel.

35. SYSTEM CLASS:

Specific systems have certain functional similarities with each other. These similarities define generic classes of systems into which each individual system falls. For example, a specific model of tank would fit into the class--Armored Vehicles.

36. SYSTEM FUNCTION:

A System Function is one of the ultimate purposes of an individual system, rather then an intermediate step leading to the performance of such a purpose.

37. SYSTEM PERFORMANCE ISSUE (SPI):

An SPI is an intermediate system action which must be performed so that its System Function can be performed. It is normally stated as a question, and it is the answer to SPI questions which must be answered in an OT. HRTES provides two formats for each SPI, normal question format and a statement format. SPI's can be written as simple actions, without modifying conditions and as actions modified by any number of specified conditions.

38. SYSTEM PERFORMANCE ISSUE CATEGORY:

SPI's in the SPI Index are divided into SPI categories. Each category has an identification number, and it is these numbers which are referenced by related System Functions.

39. SYSTEM PERFORMANCE ISSUE INDEX:

This is a collection of SPI's divided into categories.

40. TRIAL:

A trial is one performance of one Human Performance Function in one condition set by one performance unit.

41. VALUE FUNCTION:

A Value Function is a graphical representation of the values given to various possible statistic outcomes of performance of each Human Performance Function. It is based on a definition of "very good" statistic outcome, "very bad" statistic outcome, and criterion statistic outcome. Through the use of Value Functions, HRTES converts all statistic outcomes to values of those outcomes on a common scale of values.

42. WORKBOOK:

The workbook is the volume of HRTES which contains any material which is to be copied/duplicated. In general it contains: guidelines, worksheets, questionnaires, sample worksheets, and competing HRTES elements which must be rated.

operators and maintainers. Also, since human behavior is very complex, it is very difficult to determine which behavioral components must be evaluated within a total system framework. Furthermore, unlike hardware components of systems, human behvaior is highly variable; two individuals may perform in diametrically opposite fashions under identical conditions, thus producing results which are more difficult to assess.

Thus, HRTES is designed precisely to provide you with procedures which specifically address problems related to the evaluation of the human component in system evaluation. The procedures are designed to be understood easily and to ultimately yield superior system evaluation. The use of HRTES is also intended to provide other concomitant benefits. Well-designed test plans should result in the early detection of system failures and the timely incorporation of human resource considerations in system design. The application of clearly specified procedures to generate test plans should promote comparability among plans developed for similar systems. Thus, it will facilitate comparing the performance of a new system to that of an older system. This will permit easier assessment of the degree of superiority of the new system. The existence of a previous test plan developed using HRTES should reduce the level of effort required to generate the test plan for a new system and should yield savings generated through the consolidation of plans, data sharing, etc.

III. What is a System?

Before proceeding to the details of HRTES, it is important to understand a central concept--the definition of *system*. Each person or agency associated with materiel acquisition and development will have its own definition of *system* which will be legitimate to its perceived mission and needs. In the case of the materiel developer, such a definition often includes those items specified in the documents to the contractor, usually in terms of deliverable hardware and, increasingly, software items. However, the

user is concerned not only with the deliverable hardware items, but also the items to supply and maintain the *system*, and from the user's point of view, the most important item--the people to operate and maintain the hardware. As the user's representative in conducting an independent test and evaluation, you must remember that for a *system* undergoing test and evaluation, that *system* consists of: hardware and sometimes software; personnel who operate and maintain it; the training they receive; and the tools, manuals, and equipment required to use and maintain the *system*.

The definition of a system leads to a basic principle in HRTES. This principle states that "human performance is a component of total system performance." Given this principle, operator and maintainer performance is evaluated in a manner similar to the evaluation of the hardware and software components of the system. Thus, the field tester first must identify that performance which is required from the operators and maintainers if the overall system is to be able to perform satisfactorily. Then, that operator and maintainer performance must be measured within the context of the system performing its overall functions. If the measured operator or maintainer performance is shown to be inadequate, then such areas as personnel selection, training, etc. are evaluated to determine the reasons for the inadequate performance.

IV. Structure of HRTES.

HRTES consists of two logically separate phases. In the first phase you will be given a set of procedures that will result in a test plan for OT. The procedures leading to the test plan are contained in Chapters two through six of the HRTES Handbook. In the second phase you will be given procedures to evaluate tested human performance. Methods for diagnosing difficulties that have been isolated during OT have also been provided in this phase. The evaluation and diagnosis phase is covered in chapters seven and eight of the HRTES Handbook.

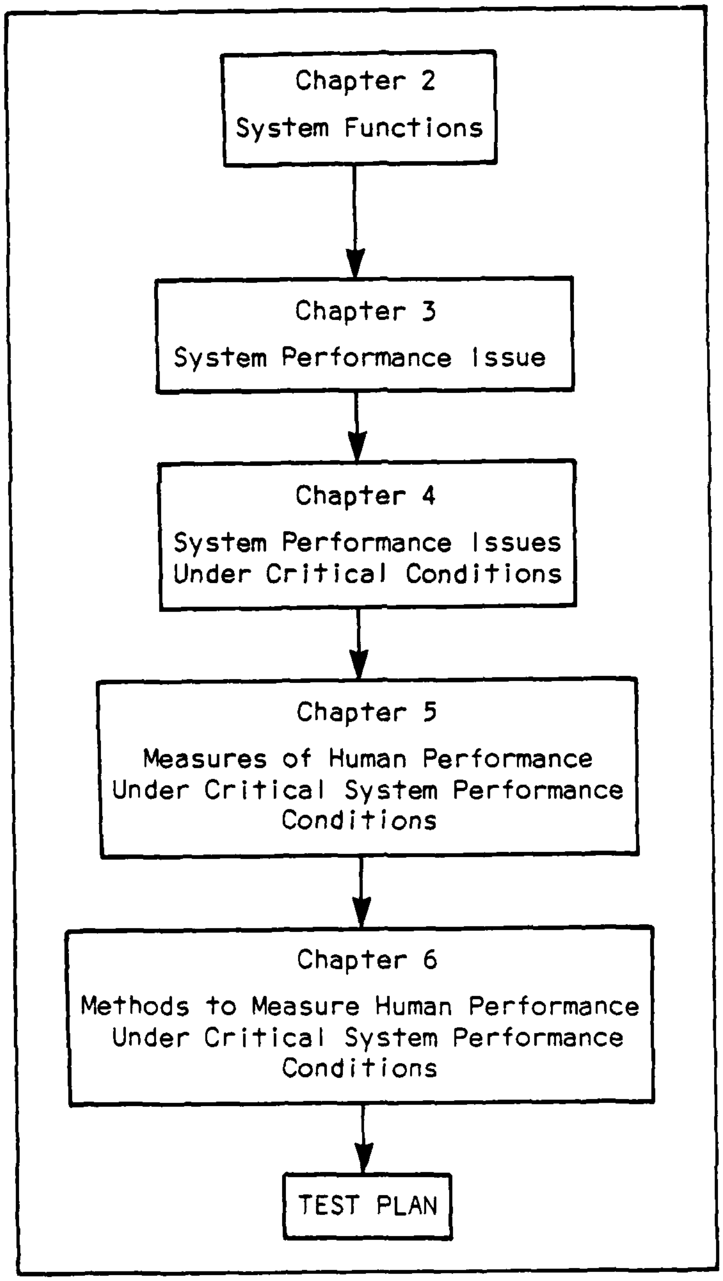
HRTES is designed to assess human performance during OT as a component of total system performance, rather than treating human performance as an isolated component. The test plan phase begins by investigating the overall functions of the system being evaluated. Later, by decomposing these global system functions into the critical system activities required to perform them, and then decomposing these activities into the critical human activities which they require, HRTES insures that the aspects of human performance which are tested are those which are critical to overall system functioning.

The test plan development phase of HRTES is schematically portrayed, by chapters of the HRTES Handbook, in Figure H1-1. This phase is designed to help you determine (1) what aspects of system operator and maintainer performance should be measured during OT, (2) which measures you should use to assess the performance, (3) what performance criteria you should use for evaluation and (4) how you should design the OT.

Test plan development begins in Chapter 2 of the Handbook by helping you identify the functions the system was intended to perform, i.e., the <u>System Functions</u>. A possible System Function of an armored system might be to "Destroy Armored Vehicles." From System Functions, HRTES moves to the procedures required to carry out the System Functions. Thus, Chapter 3 of the Handbook provides you with procedures to identify the questions that must be addressed to assess the operational effectiveness of the system. These questions are called <u>System Performance Issues (SPI's)</u>.

A possible SPI for an armored system might be "How effectively does the system acquire its targets?" Usually, a given system will have a large number of issues associated with it. HRTES provides you with a procedure to quantify the relative importance among the SPI's. The weighting scheme permits you to eliminate noncritical SPI's from the evaluation plan.

Figure H1-1
HRTES STRUCTURE (TEST PLANNING)



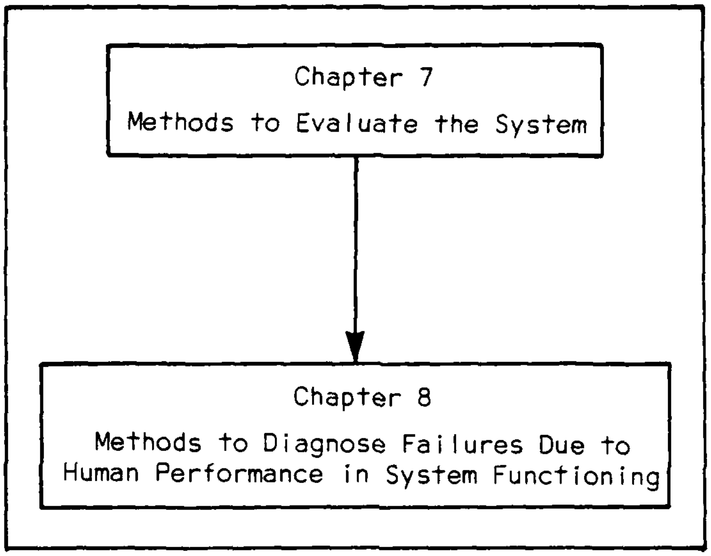
Next, Chapter 4 of the HRTES Handbook furnishes you with a method to elaborate the SPI's by adding conditions that are relevant to each SPI. For example, you might feel that precipitation may be an important condition category under which an armored system must operate. Thus, by combining a condition from this category with an SPI you might form the following question: "How effectively does the system acquire targets under blowing or falling snow?"

In Chapter 5, HRTES introduces human performance into the test plan. Here, you are given procedures to identify the specific activities, called Human Performance Functions (HPF's), that must be performed by operators and maintainers for the system to operate effectively. The HPF's are clustered into logically related categories called Human Performance Function Groups (HPF-Groups). For example, you might be interested in how well operators assemble, emplace, acquire targets, and fire an armored weapon system under blowing or falling snow. Chapter 5 also gives you a procedure to establish criteria of satisfactory human performance for various HPF's. You will be required to select the time necessary to complete HPF, the accuracy with which HPF must be completed, or a combination of time and accuracy, as a criterion to evaluate each critical HPF.

In Chapter 6 of the Handbook, HRTES gives you the means to determine the number of observations that are required during the OTs and a method to select conditions for HPF's, in addition to those conditions already attached to the SPI's. Chapter 6 also includes instructions for planning data collection during OT and instructions for collecting some diagnosis data during OT. Following completion of the field test, Chapter 6 guides you in processing the resulting data.

The second phase of HRTES, schematically portrayed in Figure H1-2, consists of evaluation and diagnosis procedures. The evaluation procedures are

Figure H1-2
HRTES STRUCTURE (EVALUATION AND DIAGNOSIS)



given in Chapter 7 of the HRTES Handbook. In Chapter 7, you will make use of the multi-attribute Evaluation Tree that the HRTES procedures required you to develop in this and in previous chapters. The Evaluation Tree is a structure relating the System, System Function, SPI's, HPF-Groups, HPF's and Statistics. Chapter 7 also offers a procedure to evaluate actual OT outcomes by using an approximate Value Functions. Finally, there are procedures that allow you to place regions of confidence around actual observed OT outcomes and procedures to determine and interpret performance values derived for various levels of the Evaluation Tree.

Chapter 8, the diagnosis chapter, offers methods to determine why various aspects of the system did not perform to expectation. First, there are procedures designed to assist you in determining which HPF's should be diagnosed. Once you have established that system failure was caused by poor human performance and not by hardware failures, HRTES provides methods to investigate whether the source of the difficulties arose because of deficiencies in (a) training, (b) Human Factors Engineering, or (c) manpower selection.

V. Using Experts.

Throughout HRTES a number of references are made to using "a group of experts" to perform one or more of the procedures in that section. At first glance, this may appear to be an implication that we have assumed that you are not fully qualified to be an operational tester. In fact, this process of using experts is a recognition that you, as an operational tester, are the person charged with the responsibility for making the technical decisions about test design and evaluation. However, the system proponent is charged with the responsibility for stating what the system is supposed to be able to do and how well it is to do it. Thus, the HRTES "group of experts" provides a mechanism for allowing you to formulate the questions to be asked in the field test, while at the same time

allowing the users to excerise their responsibility for stating system requirements.

The HRTES "group of experts" approach has an additional benefit. HRTES requires that a number of critical decisions be made throughout the test design and evaluation process. For example, it is initially important to select the most critical System Performance Issues to assure that all of the important questions will be answered by the field test. To insure that the decisions taken are acceptable to the testing, development, and user communities, the decisions must be made by the most knowledgeable individual(s) available. The most desirable strategy for making the required decisions is to have each one made by a panel of appropriate military decision makers and technical experts, including you as the test planner, who represent the thinking of the military community. If this approach is adopted, the resulting decisions should be both accurate and acceptable to various interest groups within the military community.

In most cases, the decision of who to include in the respective groups of experts is left to you. As the representative of the testing agency, you will certainly be in an advantageous position to know which experts may be needed from within your own or allied agencies. For representatives of the system propoment, it is suggested that you include personnel from the Combat Developments Center of the proponent School. You may also consider representatives of the training developer, the U.S. Army Human Engineering Laboratory, the U.S. Army Research Institute, etc.

If the strategy of using representative military decision makers and technical experts is followed, you must decide between two alternative methods for accessing and interacting with the experts. In the first approach, you prepare the appropriate HRTES materials, send them to the individual experts, receive the materials, and analyze the results.

The benefits of this method are that minimal demands are put on the experts' time and their decisions should be relatively free from confounding by discussions among the experts. However, this approach gives you little or no control over the amount of effort devoted by the experts to their tasks and a very large amount of time may be required to receive the many iterations of the HRTES process as a result of the organizational sign-off process and mail delivery time. Moreover, this approach causes problems in obtaining a consensus among the experts and in giving them further explanations to guide them through the HRTES procedures.

The second approach consists of bringing the experts together and meeting as a panel to complete the HRTES procedure(s) in question. This method negates the difficulties of the first method. It offers control over the effort of experts and their output, it requires a relatively small amount of time, it permits the achievement of a consensus, and it allows further explanations and guidance by the test planner/evaluator. Its principal disadvantage is that it requires the availability of the experts for a number of days. This second approach is strongly recommended in spite of its requirement for available time from the experts. This approach offers a higher probability of obtaining appropriate decisions, is subject to greater control by you as the test planner or evaluator, and has a higher probability of being completed within a specified time frame.

If the second alternative is adopted, you must decide how to interact with the panel of experts. Two alternatives are available. One method is for you to read through the appropriate HRTES material, make copies, distribute the copies, and make the required explanations. This alternative keeps your work load down, but it also reduces the degree to which you can guide the experts and the degree to which you can provide input. In the second alternative, you work through all the procedures which will be required of the experts prior to their meeting. With this added

experience and information, you can better understand the problems of the experts, give more complete instructions, provide more adequate leadership and detect the occurrence of errors of omission or comission. Therefore, the second alternative seems to be preferable.

In some circumstances it may prove impossible to use a group of experts, either as a panel or individually by phone and mail. In this case, it is possible for you to make all of the decisions required in HRTES. Since this puts you in the position of deciding not only the activities and conditions to be tested, but also the criteria for those activities, we strongly suggest that you use representative decision makers and experts whenever possible.

VI. How to Use HRTES.

HRTES is divided into two volumes, the Handbook and the Workbook. The Handbook consists of descriptions of the procedures which you will follow and explanations of the rationale for the procedures. The Workbook consists of those Guidelines, Worksheets, Samples, etc., which you will need to copy before they are used. By copying the appropriate pages of the Workbook you will be able to send them to other people and keep the completed Worksheets in the test file. (As used in HRTES, the word "copy" means to duplicate the page using a xerographic, electrostatic, or photographic machine copier.)

HRTES Handbook and Workbook are divided into chapters which correspond to each other. Thus, Handbook Chapter 2 relates to Workbook Chapter 2, and so forth. Normally you will use HRTES by reading sequentially each chapter of the Handbook and performing the procedures described in that chapter. Each chapter of the Handbook will direct you to copy the appropriate pages of the corresponding Workbook chapter. It should be noted that the HRTES pagination has been designed to aid you in moving from the Handbook to the Workbook. Each page number begins with either the letter H (for

Handbook), or W (for Workbook), and includes chapter and specific page designations. Thus, when you are reading chapter 5 of the Handbook and you are referred to page W5-4, you would go to Workbook chapter 5, page 4.

No individual step of the HRTES procedures is particularly complex, but the compilation of all the steps may prove somewhat difficult. It is suggested that no matter what method you choose for using HRTES, you always read carefully all the appropriate material in both the Handbook and Workbook for each procedure prior to performing it or having it performed by others. In this way you will be able to keep track of how HRTES is proceeding, understand what is about to be done, and maintain the "big picture" of your current place in the overall HRTES system.

Table H1-1 lists the general steps which you will perform with HRTES. These steps are presented here to give you an overview of all of the steps in HRTES. We realize that some of these steps may not be clear at this point. Don't worry, the following chapters of HRTES will explain each of them in detail. You can plan to refer back to this table at any time to review the relative order of these steps.

Table H1-1
STEPS IN HRTES

	STEPS	CHAPTER
1.	Select System Class(es).	2
2.	Select System Functions and begin development of Selection Tree.	2
3.	Select System Performance Issues.	3
4.	Rate conditions for each SPI and select most critical one(s).	4
5.	Combine most critical conditions with appropriate SPI's, if required, and continue develpment of Selection Tree.	4
6.	Select/combine HPF's in referenced Operational HPF-Groups and continue development of Selection Tree.	5
7.	Develop Maintenance HPF's for each SPI and continue development of Selection Tree.	5
8.	Select remaining conditions for HPF's.	5
9.	Combine conditions selected for SPI's and HPF's into condition set(s) for each HPF.	5
10.	Develop performance criterion for one trial of each HPF.	5
11.	Develop statistic criterion for multiple trials of each HPF, and continue development of Selection Tree.	5
12.	Determine data to be taken for each HPF.	6
13.	Determine number of operational units (subjects) and trials per operational unit for each HPF.	6
14.	Determine method(s) for taking data, and make appropriate preparations for taking data.	6
15.	Prepare for taking required data for diagnositc mesures.	6
16.	Process data from field test.	6
17.	Convert Selection Tree to Evaluation Tree.	7
18.	Develop Value Function for each statistic of each HPF. (Frequently, this may be done immediately following Step 11.)	7
19.	Fold back Evaluation Tree.	7
20.	Evaluate performance in OT.	7
21.	Determine which HPF's produced subcriterion performance in significant upper level node(s) of Evaluation Tree.	8
22.	Determine strategy for diagnosing inadequate HPF's.	8
23.	Take diagnostic measures and convert them to indices of Acceptability.	8
24.	Diagnose probable cause(s) of inadequate human performance.	8

VII. How to Use HRTES When You Already Have Test Plan Elements from Existing Documents.

HRTES has been designed to aid you to develop all elements required for a test planning document, without any additional information. However in some instances you may have access to earlier test planning documents that were not developed using HRTES. These existing documents may already list all or some of the following for the system to be tested: missions--equivalent to HRTES System Functions, issues and critical issues--equivalent to HRTES System Performance Issues (SPI's), factors or variables--equivalent to HRTES conditions, tasks--equivalent to HRTES criteria plus statistics, and data requirements--equivalent to HRTES data requirements. For each of these categories of test plan elements, you may be: completely satisfied and not wish any additional input from HRTES, completely dissatisfied and wish a total replacement from HRTES, or not sure and wish to accept these elements and also some equivalent ones from HRTES. Also such existing planning documents may list some categories of test plan elements while not dealing at all with other categories.

In all of these circumstances HRTES can be used to aid the test planning process. If you are completely satisfied with a given category of test plan elements from existing documents, HRTES should still be used to differentially weight these elements for the eventual diagnosis. Also HRTES should be used to relate the elements of this given category to other elements <u>from</u> which they were logically decomposed, or <u>into</u> which these elements are logically decomposed. If you are completely dissatisfied or if a given category of test plan elements is missing, you can use the appropriate chapter and section of HRTES to produce new or replacement test plan elements. If you are not sure of existing test plan elements, you can use HRTES to generate additional elements in a given category, weight both the existing and new elements, and relate both types to hierarchically higher and lower elements.

Following this page is a flow chart of how to use HRTES in these various circumstances and specifically which sections of HRTES are applicable for your situation

FLOW CHART FOR
INCLUSION OF EXISTING TEST PLAN ELEMENTS IN THE HRTES STRUCTURE

1. ALREADY HAVE MISSIONS.

YES — COMPLETELY SATISFIED? — NO/NOT SURE

(1a) RATE EXISTING MISSIONS ACCORDING TO PROCEDURES IN HRTES CHAPTER 2. DO NOT RATE HRTES SYSTEM FUNCTIONS.

(1b) IF DISSATISFIED OR NOT SURE, READ HRTES SYSTEM FUNCTIONS FOR TYPE OF SYSTEM TO BE TESTED--CHAPTERS 1 AND 2. AFTER READING, IF YOU WISH TO CONSIDER HRTES SYSTEM FUNCTIONS, TOO, THEN ADD YOUR MISSIONS TO THEM. RATE BOTH YOUR EXISTING MISSIONS AND APPROPRIATE HRTES SYSTEM FUNCTIONS ACCORDING TO PROCEDURES IN CHAPTER 2.

IF YOU HAVE INTRODUCED NEW MISSIONS INTO THE HRTES STRUCTURE, YOU WILL HAVE TO DETERMINE WHICH OF THEM DECOMPOSE INTO HRTES SYSTEM PERFORMANCE ISSUES (SPI'S). BY INSPECTION. IF YOU HAVE EXISTING ISSUES AND ARE COMPLETELY SATISFIED WITH THEM, YOU WILL HAVE TO DETERMINE WHICH NEW MISSIONS DECOMPOSE INTO THOSE ISSUES, ALSO BY INSPECTION.

2. ALREADY HAVE ISSUES.

YES — COMPLETELY SATISFIED? — NO/NOT SURE

(2a) RATE ONLY EXISTING ISSUES WHICH DEAL DIRECTLY WITH SYSTEM PERFORMANCE. IN HRTES, MAINTENANCE, SAFETY TRAINING, SELECTION, AND HFE ARE COMPONENTS OF EACH PERFORMANCE ISSUE. IF AN EXISTING DOCUMENT FORCES YOU TO KEEP THESE NON-PERFORMANCE ISSUES IN THE ISSUE FORMAT, THEN HRTES WILL PROVIDE A MEANS TO ANSWER THEM FOR THE SYSTEM AS A WHOLE (H7-20). HOWEVER, THEY SHOULD BE CONSTITUTED AS A DIFFERENT CLASS OF ISSUES--NON-PERFORMANCE ISSUES-WHICH ARE NOT RATED AND ARE HANDLED IN A COMPLETELY DIFFERENT MANNER FROM PERFORMANCE ISSUES.

(2b) IF DISSATISFIED OR NOT SURE, READ HRTES SPI'S FOR SELECTED MISSIONS TO BE TESTED--CHAPTER 3. AFTER READING, IF YOU WISH TO CONSIDER HRTES SPI'S, TOO, THEN ADD ONLY YOUR PERFORMANCE ISSUES TO THEM. DO NOT RATE OR, AT THIS TIME, INCLUDE EXISTING MAINTENANCE, SAFETY, TRAINING, SELECTION, OR HFE ISSUES. IN HRTES THESE TOPICS ARE COMPONENTS OF EACH PERFORMANCE ISSUE. IF AN EXISTING DOCUMENT FORCES YOU TO KEEP THESE NON-PERFORMANCE ISSUES IN THE ISSUE FORMAT, THE HRTES WILL PROVIDE A MEANS TO ANSWER THEM FOR THE SYSTEM AS A WHOLE (H7-20). HOWEVER, THEY SHOULD BE CONSTITUTED AS A DIFFERENT CLASS OF ISSUES--NON-PERFORMANCE ISSUES-WHICH ARE NOT RATED AND ARE HANDLED IN A COMPLETELY DIFFERENT MANNER FROM PERFORMANCE ISSUES. RATE YOUR EXISTING PERFORMANCE ISSUES AND APPROPRIATE HRTES SPI'S ACCORDING TO PROCEDURES IN CHAPTER 3.

IF YOU HAVE INTRODUCED NEW ISSUES INTO THE HRTES STRUCTURE, YOU WILL HAVE TO DETERMINE WHICH NEW ISSUES DECOMPOSE INTO APPROPRIATE GROUPS OF HRTES HUMAN PERFORMANCE FUNCTIONS (HPF'S), BY INSPECTION. IF YOU HAVE EXISTING TASK LISTS AND ARE COMPLETELY SATISFIED WITH THEM, YOU WILL HAVE TO DETERMINE WHICH NEW ISSUES DECOMPOSE INTO THOSE TASKS, ALSO BY INSPECTION.

3. ALREADY HAVE FACTORS/VARIABLES

COMPLETELY SATISFIED?

YES: (3a) DETERMINE WHICH FACTORS/VARIABLES APPLY TO WHICH SELECTED ISSUES. RATING IS NOT REQUIRED.

NO/NOT SURE: (3b) IF DISSATISFIED OR NOT SURE, READ HRTES CONDITIONS TO BE APPLIED TO SPI'S--CHAPTER 4. AFTER READING, IF YOU WISH TO CONSIDER HRTES CONDITIONS, TOO, THEN ADD YOUR FACTORS/VARIABLES TO THEM. RATE YOUR EXISTING FACTORS/VARIABLES AND APPROPRIATE HRTES CONDITIONS ACCORDING TO PROCEDURES IN CHAPTER 4.

4. ALREADY HAVE TASKS.

COMPLETELY SATISFIED?

YES: (4a) DETERMINE WHICH TASKS APPLY TO WHICH SELECTED PERFORMANCE ISSUES. THESE TASKS MUST BE APPLICABLE TO THE SYSTEM BEING TESTED. THEY MUST ALSO BE FOR BOTH OPERATIONS AND MAINTENANCE FOR EACH PERFORMANCE ISSUE. THEY MUST BE LARGE ENOUGH TO BE PRACTICAL IN A FIELD TEST, BUT THEY MUST BE SMALL ENOUGH TO PROVIDE DIAGNOSTIC UTILITY IF THEY ARE NOT PERFORMED ADEQUATELY. RATING IS NOT REQUIRED.

NO/NOT SURE: (4b) IF DISSATISFIED OR NOT SURE, READ HRTES HUMAN PERFORMANCE FUNCTIONS TO BE TESTED-CHAPTER 5. AFTER READING, IF YOU WISH TO CONSIDER HRTES HPF'S, TOO, OR IN PLACE OF THOSE YOU HAVE, THEN EITHER ADD YOUR TASKS TO THEM, OR SELECTIVELY REPLACE THE EXISTING TASKS WITH HRTES HPF'S ACCORDING TO PROCEDURES IN CHAPTER 5.

5. ALREADY HAVE CRITERIA.

CRITERIA ARE FOR?

TASKS: (5a) DETERMINE WHICH COMBINATION OF TASKS AND SETS OF CONDITIONS VARIOUS CRITERIA APPLY TO.

ISSUES: (5b) IF CRITERIA ARE FOR ISSUES ONLY, CRITERIA FOR TASKS WHICH ARE DECOMPOSED FROM THOSE ISSUES WILL HAVE TO BE DEVELOPED. USE THE ISSUE CRITERIA AS OVERALL LIMITS FOR ALL THE TASKS DECOMPOSED FROM EACH ISSUE. THEN READ THROUGH HRTES CRITERIA DEVELOPMENT--CHAPTER 5. USE THE HRTES MATERIAL TO DEVELOP TIME AND ERROR BUDGETS FOR TASKS WHICH FALL WITHIN THE OVERALL CRITERIA LIMITS FOR THE ISSUES FROM WHICH THESE TASKS ARE DECOMPOSED.

6. ALREADY HAVE DATA REQUIREMENTS FOR TASKS

COMPLETELY SATISFIED?

YES: (6a) DETERMINE WHICH TASK-CONDITION SET CRITERION PERTAINS TO EACH DATUM REQUIREMENT.

NO/NOT SURE: (6b) IF DISSATISFIED OR NOT SURE, READ HRTES OBJECTIVE AND SUBJECTIVE DATA COLLECTION--CHAPTER 6. AFTER READING, IF YOU WISH TO USE THE HRTES METHODOLOGY FOR EITHER CORRECTING UNSATISFACTORY DATA REQUIREMENTS, OR PRODUCING NEW ONES, DEVELOP OBJECTIVE AND SUBJECTIVE DATA REQUIREMENTS ACCORDING TO THE PROCEDURES IN CHAPTER 6.

EXPLANATION OF FLOW CHART

(1a) SATISFIED WITH EXISTING MISSIONS: Enter HRTES at--page H2-1. Continue through the end of the Chapter. Only copy the blank "System Function Worksheet" on page W2-14, plus associated Guidelines and Questionnaire. Record those missions with which you are satisfied on this Worksheet. Once you have rated them or received the ratings from experts you will not have to determine a cut-off point since all are to be used. You will need the ratings.

(1b) DISSATISFIED WITH EXISTING MISSIONS: Enter HRTES at--page H2-3, Section I. Continue through the end of Chapter 2. Add your existing missions from documents to the appropriate "System Function Worksheet(s)," and carry out the HRTES procedures.

(2a) SATISFIED WITH EXISTING ISSUES: Enter HRTES at--page H3-7, Section II. Continue through the end of step (10), page H3-10. Copy only your existing issues, from documents, on the "System Performance Issue Worksheet" on page W3-2. Once you have rated them or received the ratings from experts, you will not have to determine a cut-off point since all are to be used. You will need the ratings. Do not record any issues that are not performance issues, on this Worksheet. They do not need to be rated. However, make sure you retain a record of all non-performance issues (those having to do with maintenance, training, selection, HFE, etc.) so that they can be addressed in the evaluation process and answered. Read pages H7-21 to the end of the Chapter for an explanation of the reasons for dealing with non-performance issues.

(2b) DISSATISFIED WITH EXISTING ISSUES: Enter HRTES at--page H3-3, Section I. Continue through the end of Chapter 3. Add your existing issues, from documents, to the "System Performance

Issue Worksheet" on page W3-2 in addition to the HRTES SPI's to be rated. Only add existing performance issues to this Worksheet for rating. Non-performance issues should be kept track of so that they can be answered in the evaluation process, but they will be dealt with differently from performance issues in HRTES.

(3a) SATISFIED WITH EXISTING FACTORS/VARIABLES: Enter HRTES at--page H4-19, General Procedures (Final Section). Continue through the end of Chapter 4. Remember in HRTES the issues you have taken from existing documents are called System Performance Issues (SPI's) although they were not developed in HRTES.

(3b) DISSATISFIED WITH EXISTING FACTORS/VARIABLES: Enter HRTES at--page H4-3. Section I. Continue through the end of Chapter 4. Add the existing factors/variables, from documents, to the appropriate "Condition Rating Worksheets (method 1 or 2) depending on their type. If some factors/variables are redundant to listed HRTES conditions, do not record them on the Worksheet.

(4a) SATISFIED WITH EXISTING TASKS: Enter HRTES at--page H5-20, Section IV. Continue through the end of Chapter 5, skipping step (3). In HRTES, tasks are called Human Performance Functions (HPF's), and functionally related collections of tasks which are performed continuously are called HPF-Groups.

(4b) DISSATISFIED WITH EXISTING TASKS: Enter HRTES at--page H5-4, Section I. Continue through the end of Section II. Then re-enter at page H5-20, Section VI, and continue through the end of Chapter 5, skipping step (3). You will have to determine which HPF-Groups to examine. If you have used HRTES SPI's, they will have referred you to appropriate HPF-Groups.

Otherwise you will have to look through the various "Operational HPF-Group Worksheets" and decide which ones fit your needs. Once you have determined this, add your existing tasks as appropriate, and follow the HRTES procedures. HRTES does not contain completed HPF-Group Worksheets for maintenance, due to the hardware operation of maintenance. It does contain a blank "Maintenance Worksheet" on page W5-48 and associated Guidelines. Depending upon your level of dissatisfaction with your existing maintenance tasks, it may be reasonable to insert them in this Worksheet. However if you are completely dissatisfied with them, and you have sufficient time, you may find it reasonable to simply eliminate these maintenance tasks and follow the HRTES maintenance HPF procedure without any additions.

(5a) SATISFIED WITH CRITERIA FOR ISSUES. HAVE NO CRITERIA FOR TASKS: Enter HRTES at--page H5-20, Section VI, and make sure to complete steps (3) and (4) of this procedure. If you have acceptable issue criteria, the problem of defining task (HPF) criteria of time and accuracy will be greatly aided since these task (HPF) criteria are components of the issue criteria and must add up to those issue criteria. Therefore if you have such issue criteria, but still must develop HPF criteria, append "Guidelines for Assigning HPF's Criteria and Statistics Within Their Overall SPI Criterion (PW1-3)" for both "Guidelines for Developing Performance Criteria" and "Guidelines for Developing Statistics and Statistic Criterion" pages W5-51 through W5-63. Copy the name of the SPI. Enter an "X" in the appropriate box to indicate whether the SPI criterion, and therefore the component HPF's apply to operations or to maintenance required for those operations. Then list all those HPF's (operational or maintenance) that are components of the SPI to be tested and which will require separate, individual criteria and statistics.

(5b) SATISFIED WITH CRITERIA FOR TASKS: Enter HRTES at--page W5-20, Section VI, and make sure to complete steps (3) and (4) of this procedure. You will have to be quite careful to determine to which combination of task (HPF) plus variables (conditions a given criterion belongs, so that it applies to the correct node of the Selection Tree. You should also make sure to differentiate issue from task criteria. If you do not have any task criteria, it will be difficult to determine the causes of inadequate issue performance later.

(6a) SATISFIED WITH DATA REQUIREMENTS FOR TASKS: If you are satisfied with all aspects of data requirements for tasks (that is objective data requirements, subjective data requirements, and number of trials of each task (HPF) to insure reliability) enter HRTES at--page H6-11, What Data to Collect. Continue through step (5) on page H6-13. Simply record the data to be taken on the HRTES "Performance Data Collection Worksheet", page W6-2, for each separate criterion of each HPF (task) that has been produced. These Worksheets are to be sent to the field test personnel for recording performance data. Then re-enter HRTES at page--H6-15, Section III, and complete the procedures for Training Data on page H6-17. This training data is to be sent to test trainers prior to the opening of the training phase of the field test.

(6b) DISSATISFIED WITH SOME ASPECTS OF DATA REQUIREMENTS FOR TASKS: Due to the complexity of this chapter of HRTES, this question is presented in the following tabular form.

DISSATISFIED WITH:	ENTER HRTES:	CONTINUE TO:
DETERMINING NUMBER OF TRIALS FOR EACH HPF (TASK)	PAGE H6-3, SECTION I.	END OF SECTION I ON PAGE H6-10.
OBJECTIVE DATA TO COLLECT FOR EACH HPF (TASK)	PAGE H6-11, WHAT DATA TO COLLECT	END OF SECTION II ON PAGE H6-15
SUBJECTIVE DATA TO COLLECT FOR EACH HPF (TASK)	PAGE H6-15, SECTION III	END OF STEP (7), PAGE H6-18

2. SYSTEM INDEX

CONTENTS:
- I. Define the System to be Tested.
- II. System Index.
- III. Obtaining Rating of System Functions.
- IV. Filling Out Selection Tree.

ACTIONS:
- (1) Selection of System Class(es) in which your system fits.
- (2) Rating criticality of System Functions.
- (3) Selection of critical System Functions based on ratings.
- (4) Initial preparation of Selection Tree and inclusion of selected System Functions and their weights.

PRODUCTS:
- (1) Completed System Function Worksheet.
- (2) Selection Tree containing System Functions and their weights.

USED FOR:
- (1) Working Papers for preparation of Independent Evaluation Plan.

Developing Appropriate Test Elements for This Chapter

To avoid developing test elements (System Functions, System Performance Issues, conditions, Human Performance Functions) in this chapter, when they should be developed in a later chapter, look through the following list of examples which are related. They are not complete, but they should give you an idea of the sorts of testing elements to be developed now and those which will be developed in later chapters.

YOU ARE HERE

SYSTEM FUNCTION EXAMPLE: (CHAPTER 2)

Destroy enemy units (aircraft, armored vehicles, fixed emplacements, personnel, etc.).

SYSTEM PERFORMANCE ISSUE EXAMPLES: (CHAPTER 3)

Ability to be transported.

Maneuver in travel.

Navigation

Target acquisition.

Maneuver in attack/defense

Establishment and maintenance of communications.

Delivery of ammunition on target.

CONDITION CATEGORY EXAMPLES: (CHAPTER 4)

Illumination.

Ground slope.

Ground surface.

Target characteristics (type, number speed, location, direction of motion, etc.

Tactical characteristics (number of friendly systems employed, speed, location, direction of motion, concealment, etc.).

Protective gear worn.

Time since end of training.

Duration of preceeding work.

HUMAN PERFORMANCE FUNCATION EXAMPLES: (CHAPTER 5)

Load system.

Unload system.

Read vehicle maneuvering instruments.

Travel designated route.

Detect target(s).

Identify targets(s).

Prioritize targets.

Perform tight turn-forwards.

Establish communications net.

Load ammunition.

Aim weapon.

Fire weapon.

Reorient system to next target.

1. Defining the System to be Tested.

For each major group of HRTES procedures that follow, you will decide whether to rely on one or more experts, or to perform the procedure entirely yourself. If you decide to use experts other than yourself to perform a group of procedures, you may have the problem of different experts having varying definitions of the nature of the system being tested. For example, some experts may define an air defense weapon system to include only the weapon delivery and target acquisition components. Other experts may include the command and control, transportation, and resupply components. To alleviate this problem, HRTES includes a "Definition of System to be Tested" worksheet, page W2-2. Use the following procedure to complete this worksheet.

(1) Copy the "Definition of System to be Tested" worksheet, page W2-2.

(2) List the full name of the system. Notice that the name may change slightly with subsequent tests. For example, as shown on the Sample Worksheet on page H2-5, the name includes only the air defense launcher for the OT I; whereas, a full air defense battery is name for the OT II, as shown on the Sample Worksheet on page H2-6.

(3) List all items that constitute <u>one</u> system to be tested. As with the name of the tested system, this will probably vary with the test being conducted. Notice, for example, the differences in items listed on the Sample Worksheets for the MERCURY Air Defense System.

(4) List all of the operator and maintenance personnel who are required for the system. The number and type of people will

largely depend upon the number and type of items identified in step (3) above. (Do not list any replacement or duplicate personnel that may be used as "player personnel" in the test - only list those people necessary for one full crew.)

(5) List the name of the system (or systems) that would be either replaced or augmented by the system being tested. The system(s) being replaced or augmented may be radically different in configuration from the new system, but its functions will be similar. For example, a seige cannon replaced a catapult. Listing the system to be replaced or augmented may help the experts to identify the functions and issues associated with the new system.

(6) If useful, draw a diagram of the system. If a single system consists of a number of geographically dispersed items, this diagram may be helpful in reminding the experts of some of the functions and issues of the new system. If only one, or few, items are to be included in the definition of the system, this diagram may be omitted.

(7) Include a copy of the completed "Definition of System to be Tested" worksheet whenever you give any forms or procedures to other experts.

SAMPLE

HRTES DEFINITION OF SYSTEM TO BE TESTED

NAME OF TESTED SYSTEM:

MERCURY Air Defense Missile Launcher

ITEMS TO BE INCLUDED IN TEST (Number and Type):

1 - MERCURY Self Propelled Air Defense Launcher
(including: full missile load, and
Ballistics Computer System)

OPERATOR AND MAINTAINER PERSONNEL (Number and Type):

1 - Loader

1 - Driver

1 - Launcher Commander

No maintenance personnel to be included. Field maintenance to be performed by driver and loader.

SYSTEM TO BE AUGMENTED OR REPLACED:

Artemis Short Range Air Defense System and
Blunderbus Air Defense Gun

DIAGRAM OF SYSTEM TO BE TESTED:

Not Applicable

SYSTEM MERCURY Air Defense System TEST OT I DATE 5 Jun 79 PAGE

NAME ______ TELEPHONE ______

SAMPLE

HRTES DEFINITION OF SYSTEM TO BE TESTED

NAME OF TESTED SYSTEM:

MERCURY Air Defense Battery

ITEMS TO BE INCLUDED IN TEST (Number and Type):

3 - MERCURY Self Propelled Air Defense Missile Launchers (including: full missile load, and *Ballistics Computer System*)

1 - Flycatcher Radar System

1 - Norbert Weiner Self Propelled C^3I Unit

3 - Improved Tortoise Tracked Vehicles (1 - towing Flycatcher Radar, 1 - resupply vehicle, and 1 - maintenance vehicle)

OPERATOR AND MAINTAINER PERSONNEL (Number and Type):

3 - Loaders (1 per missile launcher)

7 - Drivers (1 per missile launcher, 1 per Tortoise vehicle, 1 for C^3I vehicle)

3 - Launcher Commanders (1 per missile launcher)

1 - Radar Operator

1 - C^3I System Operator

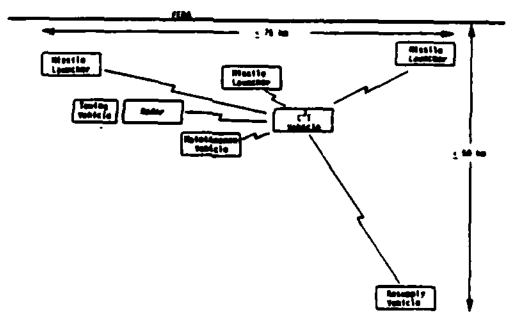
1 - Unit Commander (operates from C^3I unit)

3 - Maintenance Technicians (1 - electronics, 1 - missile, 1 - vehicle)

SYSTEM TO BE AUGMENTED OR REPLACED:

Artemis Short Range Air Defense System and Blunderbus Air Defense Gun

DIAGRAM OF SYSTEM TO BE TESTED:



SYSTEM MERCURY Air Defense System TEST OT II DATE 25 Mar 81 PAGE

NAME ______ TELEPHONE ______

II. System Index.

The System Index on the following page is a list of System Classes. This list is designed to help you determine in which class the system being tested belongs. Once you have read through the list and decided where your system belongs, write down the reference HRTES Workbook page number. If your system belongs in more than one System Class, write down the Workbook page number for each additional System Class. The pages in the Workbook to which you are referred consist of "System Function Worksheets." Each Worksheet contains System Functions which are relevant to the System Class you selected.

III. Obtaining Ratings of System Functions.

System Functions are the purposes of a system, not intermediate steps leading to those purposes. Normally they are purposes which can be performed by an individual system. To develop a test plan, it is necessary that the most critical System Functions be identified and included. To insure that the field test is not impractically long or expensive it is necessary that less critical System Functions be identified and eliminated from consideration. To evaluate field test results, it is necessary that the relative criticality of System Functions, which are tested, be specified. The process of weighting the criticality of System Functions, leading to inclusion in the test plan, is described below. The ratings may be made by a group of experts including you, or by you independently (see introduction, page H1-9).

(1) Refer to the Workbook page number(s) you wrote down in part I to find the "System Function Worksheet" for your System Class in the HRTES Workbook.

(2) Copy the "System Function Worksheet," and add any System Functions you find relevant to your System Class which are not present. These new System Functions will be treated in the same manner as those already listed.

SYSTEM INDEX

System Class	Workbook Page
1. Air Defense Weapons, including: missiles, guns and high energy systems.	W2-3
2. Armored Vehicles, including: battle tanks, fighting vehicles, reconnaissance vehicles, armored personnel carriers, and anti-armor weapons (mounted).	W2-4
3. Aviation Systems, including: helicopters and fixed-wing aircraft.	W2-5
4. Battlefield Communication Systems, including: man-portable radios, vehicle-portable radios, visual communications systems, and base radio systems.	W2-6
5. C^2 (C^2I) Systems, including: fire control systems.	W2-7
6. Combat/Tactical Support Equipment, including: combat engineer vehicles, recovery vehicles, demolition equipment, and bridging equipment.	W2-8
7. Electronic Warfare and Surveillance Systems, including: countermeasures equipment and sighting and surveillance equipment.	W2-9
8. Ground Transportation Equipment, including: utility trucks, medium trucks, and heavy trucks.	W2-10
9. Infantry Weapons, including: point target weapons, and area weapons, man-portable anti-armor weapons, and man-portable anti-aircraft weapons.	W2-11
10. Ordance Systems, including: tube artillery and missile artillery.	W2-12
11. Target Acquisition and/or Designator Systems.	W2-13

(3) Make as many copies of the "System Function Worksheet(s)" as there are experts to do the rating.

(4) Copy "Guidelines for Selecting System Functions" and the "Sample System Function Worksheet" on page W2-14 and page W2-16 for each expert.

(5) Copy the "System Function Rating Questionnaire" on page W2-17, for each expert.

(6) Submit the above copies to the appropriate experts. Check that your submission includes:

(a) System Function Worksheet(s).
(b) Guidelines for Selecting System Functions.
(c) System Function Questionnaire.
(d) Sample Worksheet.

(7) Upon receipt of the completed "System Function Worksheets" from all experts, you will aggregate their ratings. Do so by computing the means of the Sum of Ratings for each System Function and list them in their appropriate places in the "Sum of Ratings" column of a fresh Worksheet. (See sample "System Function Worksheet," page H2-10). Aggregation may also be accomplished by discussion with, and consensus of the experts. The former procedure will insure independence of the data obtained, and will tend to reduce the time required for its production. The latter procedure will tend to produce data with a high degree of acceptability to the experts, and might produce superior data throught the process of discussion.

SAMPLE

HRTES

SYSTEM FUNCTION WORKSHEET FOR AIR DEFENSE WEAPONS

SYSTEM CLASS:
AIR DEFENSE WEAPONS including: Short Range Missiles, Medium Range Missiles, Long Range Missiles, Air Defense Guns, High Energy Systems

SYSTEM FUNCTIONS	SYSTEM FUNCTION RATINGS: 1. DESIGNED FOR SYSTEM FUNCTION	2. PERFORMED WITH OTHER FUNCTIONS	3. PROBABILITY OF SYSTEM LOSS	4. PROBABILITY OF UNSUCCESSFUL CONDUCT	5. GENERIC SYSTEMS HAD PROBLEMS	6. SPECIFIC SYSTEM HAD PROBLEMS	SUM OF RATINGS/ SELECTION WEIGHT	SPI CATEGORY REFERENCES
1. Destroy aircraft.							24	1,2,4 5,7
2. Confuse and disrupt aircraft.							10	1,2,4 5,7
3. Deny selected airspace/formation to attacking aircraft.							20	1,2,4 5,7
4. Destroy ground targets.							1	1,2,5 7
5. Protect operator/crew from enemy action.							17	2,3,7

36	
35	
34	
33	
32	
31	
30	
29	
28	
27	
26	
25	
24	1
23	
22	
21	
20	3
19	
18	
17	5
16	
15	
14	
13	
12	
11	
10	2
9	
8	
7	
6	
5	
4	
3	
2	
1	4
0	

SYSTEM *MERCURY AIR DEFENSE WEAPON SYSTEM* TEST *OT II* DATE *25 MAR 81* PAGE ____

NAME ____________________ TELEPHONE ____________

(8) Record the System Function identification numbers in the last column according to the mean Sum of Ratings (e.g., number 1 would be written beside "24" in the column if System Function 1 had a mean sum of ratings of 24.) This column gives you a graphic aid for selecting the cut-off point.

(9) Select the System Function(s) which must definitely be tested, based on the aggregated ratings, by establishing a cut-off point in the last column of the Worksheet.

This procedure may be done by you independently, but it is recommended that it be done jointly with all the experts. If you cannot consult experts directly, take into consideration the cut-off points which they established on their Worksheets. The process of selecting the cut-off point should include consideration of: costs of testing differences between ratings; relationships between System Functions and previous requirements listed in the ROC, LOA and/or MENS.

IV. Filling Out the Selection Tree.

The Selection Tree serves as an ongoing record of items you have selected and rated, and it also indicates the relationships among these items. Careful preparation of the Selection Tree throughout each stage of this Handbook will aid you in understanding the selection process and <u>will be absolutely necessary for the OT evaluation process</u>. For this reason, a copy of it must be included with OT plans so that evaluation personnel will have access to it.

SOME TECHNICAL SUGGESTIONS FOR CONSTRUCTING THE SELECTION TREE:

A tree structure consists of nodes and branches. The root of the tree is a node from which everything starts. (Usually the root of such a tree is drawn at the top.) The root node branches into several other nodes. Each node branches into more new nodes, and so on. Usually a tree structure is used to represent and simplify some existing relationship in "real life." In HRTES, we use tree structures to aid in the process of selecting the elements to be tested, and then to evaluate the results.

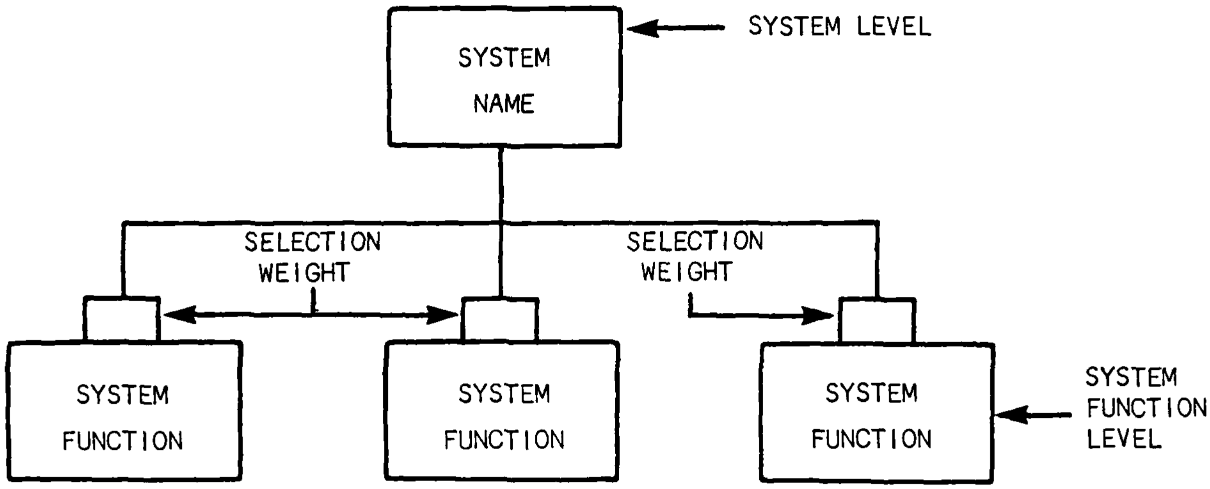
A tree structure, by its nature, has the tendency to expand considerably and, in many cases, it is hard to include a whole tree on one sheet of paper. Since we are unable to predict the exact structure of the tree for each system to be tested (i.e., to know how many System Functions will be chosen, and then for each System Function how many SPI's will be chosen, etc.), it is impossible to supply you with a given tree structure here.

The process of building the Selection Tree and filling out the necessary information is parallel to the process of selecting the right elements for testing. This process is done by you. HRTES will guide you in drawing your own Selection Tree, and filling it in, level by level, as you are working through the HRTES chapters.

At this point, start the process of building the Selection Tree.

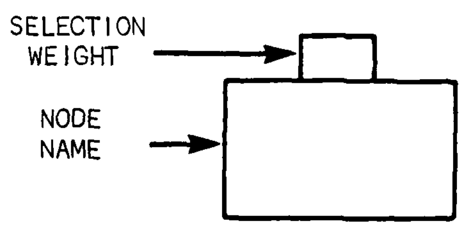
You have developed information for structuring the first two levels: The System Level (which is the root of this tree), and the System Function Level (which consists of nodes branching from the System Node).

For example, the structure of the first two levels of a Selection Tree when 3 System Functions were selected is given below:



(1) Choose a large sheet of paper on which you will draw the *Selection Tree*.

(2) Draw the structure of the first two levels of the Selection Tree in the center of the top of the page. Draw one node for the System and in the second level draw nodes for each System Function chosen. The suggested node format is as follows:



This structure of a node enables you to fill in the name of each node and its selection weight.

(3) Write the name of the System in the System Node and the names of System Functions chosen in the System Function Nodes.

(4) Fill in the selection weight for each System Function.

The selection weight of a System Function is the "Sum of Ratings" for the System Function which you recorded on the fresh "System Function Worksheet."

(5) Keep your Selection Tree. You will need it throughout the OT.

3. SYSTEM PERFORMANCE ISSUE (SPI) INDEX

CONTENTS:	I.	Selecting System Performance Issues.
	II.	Obtaining Rating of System Performance Issues (SPI's).
ACTIONS:	(1)	Examination of SPI's within referenced SPI Categories.
	(2)	Rating of the criticality of SPI's for each System Function.
	(3)	Selection of critical SPI's based on ratings.
PRODUCTS:	(1)	Completed System Performance Issue Worksheet(s).
	(2)	Completed SPI Cutoff Worksheet(s).
USED FOR:	(1)	Working Papers for further HRTES processes.

Developing Appropriate Test Elements for This Chapter

To avoid developing test elements (System Functions, System Performance Issues, conditions, Human Performance Functions) in this chapter, when they should be developed in a later chapter, look through the following list of examples which are related. They are not complete, but they should give you an idea of the sorts of testing elements to be developed now and those which will be developed in later chapters.

SYSTEM FUNCTION EXAMPLE: (CHAPTER 2)

Destroy enemy units (aircraft, armored vehicles, fixed emplacements, personnel, etc.).

YOU ARE HERE

SYSTEM PERFORMANCE ISSUE EXAMPLES: (CHAPTER 3)

Ability to be transported.
Maneuver in travel.
Navigation.
Target acquisition.
Maneuver in attack/defense.
Establishment and maintenance of communications.
Delivery of ammunition on target.

CONDITION CATEGORY EXAMPLES: (CHAPTER 4)

Illumination.
Ground slope.
Ground surface.
Target characteristics (type, number speed, location, direction of motion, etc.
Tactical characteristics (number of friendly systems employed, speed, location, direction of motion, concealment, etc.).
Protective gear worn.
Time since end of training.
Duration of preceeding work.

HUMAN PERFORMANCE FUNCTION EXAMPLES: (CHAPTER 5)

Load system.
Unload system.
Read vehicle maneuvering instruments.
Travel designated route.
Detect target(s).
Identify target(s).
Prioritize targets.
Perform tight turn-forwards.
Establish communications net.
Load ammunition.
Aim weapon.
Fire weapon.
Reorient system to next target.

I. Selecting System Performance Issues.

System Performance Issues (SPI's) are the intermediate actions which the system must take to perform each System Function. Normally, they are formatted as questions. It is the answers to these SPI questions which are sought in the field test. System Performance Issues are also used to help clarify the Operational Issues and the Critical Test Issues that are included in the Independent Evaluation Plan. In the "System Performance Issue Index," SPI's are presented in a two part format: the actual SPI question, and a preceding SPI statement. The SPI statement portion has been added for two reasons: it is easier to rate a statement than a question, and a statement takes less space in various forms than a question. When SPI's are ultimately included in test plans, they should revert to their question format.

The "System Function Worksheet" from Chapter 2 referred you to the relevant SPI Categories (not to the specific SPI's) for each System Function that you selected. Find those SPI Categories in the "System Performance Issue Index," found on the following pages. You will select those SPI's which you consider must be included in the OT. You may also select other SPI's found in the Index, even though they were not specifically referenced. Notice that the SPI's are stated as general actions, without limiting qualification. Thus, no mention is made of such things as tactics, training, etc. These qualifications, known as "Conditions" will be considered later in Chapter 4.

The "System Performance Issue Index" refers you to "Operational HPF-Group(s)." There are Operational and Maintenance HPF-Groups. The procedure for developing Maintenance HPF-Groups will be given in Chapter 5.

SYSTEM PERFORMANCE ISSUE INDEX

SPI Category	SPI Statement	SPI Question	Operational HPF-Group References
1. Weapon Delivery	Target acquisition	How effectively can the system acquire its targets?	2
	Delivery of ammunition on target	How effectively can the system deliver its ammunition on the the target/target area?	1 - Weapon Delivery 3 - Grnd to Grnd Missiles 4 - Hand Grenades 5 - Mines
	Engagement of several targets	How effectively can the system engage several targets, simultaneously?	1 - Weapon Delivery 3 - Grnd to Grnd Missiles 4 - Hand Grenades
2. Maneuverability	Navigation	How effectively can the system navigate?	27
	Maneuver in travel	How effectively can the system maneuver in travel?	24 - Grnd Vehicles 25 - Helicopters
	Maneuver in attack/ defense	How effectively can the system maneuver in attack/defense?	24 - Grnd Vehicles 26 - Helicopters
	Self-Recovery	How effectively can the system engage in self-recovery?	30
3. Vulnerability/ Survivability	Prevention of detection/location	How effectively can the system prevent its detection and accurate location?	24 - Grnd Vehicles 26 - Helicopters
	Escape from system	How effectively can the operator(s)/troops escape from the system?	31
	Protection of operator(s), etc.	How adequately can the system protect its operator(s)/ troops/materiel from small arms fire and minimize the effects of major weapon fire?	32
	Movement of system, between operations, to prevent location	How effectively can the system be moved, between operations, to minimize the probability of detection/location?	28 25 - Grnd Vehicles 26 - Helicopters
4. Command and Control	Representation of battlefield conditions	How effectively does the system represent terrain/obstacles/ installations/weather?	10
	Representation of status of forces	How effectively does the system represent the status of forces?	11

SYSTEM PERFORMANCE ISSUE INDEX (Continued)

SPI Category	SPI Statement	SPI Question	Operational HPF-Group References
Command and Control (continued)	Projection of battle-field operations	How effectively does the system project battlefield operations?	12
	Projection of weather conditions	How effectively does the system project weather conditions?	13
	Selection and ordering of targets	How effectively does the system select and order targets for attack?	15
	Management of weapon functions	How effectively does the system manage weapon functions?	9
	Personnel planning	How effectively does the system prepare personnel plans?	18
	Logistics recommen-dations	How effectively does the system recommend logistics procedures?	17
	Selection of Friendly forces	How effectively does the system select the most appropriate friendly forces to engage in an operation?	14
	Battlefield control of friendly forces	How effectively does the system control friendly forces on the battlefield?	16
5. Communications	Establishment and maintenance of communications	How effectively does the system establish/maintain communi-cations between organizational nodes?	21
	Prevention of inter-ception/jamming	How effectively does the system prevent interception/jamming of its communications?	22
	Information routing	How effectively does the system identify and route output to the most appropriate nodes of the organization?	23
6. Reconnaissance	Information gathering	How effectively does the system gather appropriate information?	20
	Fire control-recon-naissance	How effectively does the system engage in fire control?	19
7. Transportation	Ability to be trans-ported	How effectively can the system be transported?	24 - Grnd Vehicles 25 - Helicopters

SYSTEM PERFORMANCE ISSUE INDEX
(Continued)

SPI Category	SPI Statement	SPI Question	Operational HPF-Group References
Transportation (continued)	Delivery of cargo	How effectively can the system deliver the troops/materiel in fully operable condition?	24 - Grnd Vehicles 25 - Helicopters 26 - Helicopter (tactical)
	Loading/Unloading	How effectively can the system be loaded/unloaded with troops/ applicable materiel/fuel/ ammunition/wounded personnel?	28
8. Target Acquisition and Designation (Performed by Independent Acquisition and/ or Designation systems)	Acquistion of targets	How effectively does the system acquire its targets?	2
	Target information gathering and interpretation	How effectively does the system gather the appropriate information about the targets and interpret that information into meaningful data?	6
	Target behavior prediction	How effectively does the system predict target behavior?	8
	Delivery of designator on target	How effectively does the system designate the appropriate targets?	7
9. Engineering	Vehicle recovery	How effectively can the system recover a disabled vehicle?	29
	Obstacle removal	How effectively can the system remove/breach obstacles?	34
	Bridging	How effectively can the system bridge an obstacle?	35

II. Obtaining Ratings of System Performance Issues.

The purpose in rating SPI's is to aid you in determining which SPI's to include in the Operational Test. Such SPI's can be modified, or added to, by the inclusion of conditions that you consider important. The procedure for modifying SPI's will be discussed fully in Chapter 4. The ratings of SPI's may be made by a group of experts including you, or by you independently (see Introduction page H1-9).

Selection of SPI's consists of two stages. In the first stage, the selection weight of each SPI is obtained according to given attributes of criticality. At this stage you may be assisted by experts. In the second stage, you will basically be doing technical computation in which you will use the selection weights of SPI's and the selection weights of System Functions. The selection of SPI's will be based on these computations.

Instructions (1) through (7) are stage 1 of the process.

(1) Refer to page W3-2 , and copy one "System Performance Issue Worksheet" for each System Function that has been selected.

(2) Enter the name of one System Function per "SPI Worksheet," along with the System Function selection weight you computed for it in the previous chapter. (See "Sample System Performance Issue Worksheet," pages W3-5 and W3-6 .)

(3) Each System Function referred you to some SPI Categories (see the "System Function Worksheet" you completed in Chapter 2). Write in the column titled "SPI's for this System Function" those SPI statements contained within the SPI Categories that were referenced by this System Function

(the SPI's are listed in the SPI Index). If you think a whole SPI Category is irrelevant to your system, you may exclude the Category although it was referenced. In the case of new System Functions, no reference to SPI's will exist, and you will have to decide which SPI's are applicable yourself.

(4) In the "SPI Index" in this chapter, each SPI refers to one or more HPF-Groups. For each SPI, write the HPF-Group reference number in the "HPF-Group Reference Column." Usually, an HPF-Group is general and will apply to several classes of systems. In some cases an SPI will refer to several HPF-Groups, each referring to a different system class. In each case, select the HPF-Group that applies to your system.

(5) Add any other SPI's which you think may be appropriate for the System Function, but which are not present. These new SPI's will be treated in the same manner as those already listed.

(6) For each expert:

(a) Make copies of the "System Performance Issue Worksheets" you have prepared.
(b) Copy "Guidelines for Selecting SPI's" on page W3-3 and the 2-page, "Sample Worksheet" on pages W3-5 and W3-6.
(c) Copy the "SPI Rating Questionnaire" on page W3-7 .

(7) Submit the above package of worksheets and guidelines to the appropriate experts. Check that you submission includes:

(a) Guidelines for Selecting SPI's and Sample Worksheets.

(b) SPI Rating Questionnaires.

(c) System Performance Issue Worksheets.

(Instructions (8) through (14) are stage 2 of the process.)

(8) Upon receipt of the completed "System Performance Issue Worksheets," you are to aggregate the results from all experts. Compute the mean of each SPI selection weight, and record it on a fresh Worksheet in the "SPI Selection Weight Column." We have supplied you with a sample "SPI Worksheet" (pages H3-12 and H3-13) which shows a sample of <u>mean</u> ratings and other computations that you will have to make. Aggregation may also be accomplished by discussion with and consensus of the experts, rather than computing formal means.

(9) Convert all the selection weights of SPI's related to the same System Function to <u>relative weights</u>. (This procedure simply normalizes the selection weights into numbers between 0 and 1 which sum to 1.) To perform this procedure, for each System Function do the following:

(a) Sum all the selection weights of SPI's of a given System Function, and record the result in the "Grand Total" box in the Worksheet.

(b) Divide each SPI's selection weight by the grand total you computed in (a). The result is called <u>Relative Weight of the SPI</u>.

(c) Record the relative weight of each SPI in the appropriate column.

(10) Multiply the relative weight of each SPI by the selection weight of its System Function. Record the result in the "Product" column on the Worksheet.

(11) Select and mark the SPI with the highest product for each System Function by putting an asterisk (*) in the "Selected SPI (*)" column. If you have chosen only one relevant System Function to evaluate, you will have only one selected SPI at this point. (You will add additional asterisks in steps (13) and (14).)

(12) Refer to page W3-9 and copy the "SPI Cutoff Worksheet." On this Worksheet list <u>all your unselected SPI's together</u> in order of their products (from highest to lowest), without regard to their originating System Functions.

Though these SPI's are not organized by their System Function, you must note in the "Related System Function" column the System Function to which each SPI belongs for later use. (See "Sample SPI Cutoff Worksheet," H3-14).

(13) Start at the top of this list and select any additional SPI's you wish to include in the OT by establishing a numerical cut-off point. SPI's below this product should be excluded, SPI's above should be included. This procedure may be done by you independently, but it is recommended that it be done jointly with the experts.

The process of selecting the cut-off point should include consideration of: costs of testing; differences in ratings between adjacent SPI's; relationship between SPI's and previous requirements listed in the MENS, ROC, and LOA.

(14) Go back to the "System Performance Issue Worksheets" and enter an asterisk for these additional selected SPI's in the "Selected SPI (*)" columns of the "SPI Worksheets." (See "Sample SPI Worksheet," pages H3-12 and H3-13.)

SAMPLE

HRTES

SYSTEM PERFORMANCE ISSUE WORKSHEET

SYSTEM FUNCTION: DESTROY AIRCRAFT

SYSTEM FUNCTION SELECTION WEIGHT: 24

SPI's FOR THIS SYSTEM FUNCTION	HPF-GROUP REFERENCE	SPI RATINGS: 1. PERFORMED WITH OTHER SPI's	2. PROBABILITY OF SYSTEM FUNCTION FAILURE	3. GENERIC SYSTEMS HAD PROBLEMS	4. SPECIFIC SYSTEMS HAD PROBLEMS	SUM OF RATINGS/ SELECTION WEIGHT	RELATIVE WEIGHT	PRODUCT	SELECTED SPI (*)
TARGET ACQUISITION	2					23.5	.15	3.6	*
DELIVER AMMUNITION ON TARGET	1					19.0	.12	2.9	*
ENGAGE SEVERAL TARGETS	1					18.5	.12	2.9	*
NAVIGATION	27					6.0	.04	1.0	
MANEUVER IN TRAVEL	24					10.5	.07	1.7	
MANEUVER IN ATTACK / DEFENSE	24					1.5	.01	0.2	
SELF-RECOVERY	30					7.0	.05	1.2	
ESTABLISHMENT & MAINTENANCE OF COMMUNICATIONS	21					9.5	.06	1.4	
PREVENTION OF INTERCEPTION/JAMMING	22					21.5	.14	3.4	*
INFORMATION ROUTING	23					6.0	.04	1.0	
ABILITY TO BE TRANSPORTED	24					0	0	0	
DELIVERY OF CARGO	24					19.5	.13	3.1	*
					GRAND TOTAL	CONT.			

SYSTEM MERCURY AIR DEFENSE WEAPON SYSTEM TEST OT-II DATE 25 MAR 81 PAGE ____

NAME ____________________ TELEPHONE ____________

SAMPLE

HRTES

SYSTEM PERFORMANCE ISSUE WORKSHEET

SYSTEM FUNCTION: DESTROY AIRCRAFT

SYSTEM FUNCTION SELECTION WEIGHT: 24

SPI's FOR THIS SYSTEM FUNCTION	HPF-GROUP REFERENCE	SPI RATINGS: 1. PERFORMED WITH OTHER SPI's	2. PROBABILITY OF SYSTEM FUNCTION FAILURE	3. GENERIC SYSTEMS HAD PROBLEMS	4. SPECIFIC SYSTEMS HAD PROBLEMS	SUM OF RATINGS/ SELECTION WEIGHT	RELATIVE WEIGHT	PRODUCT	SELECTED SPI (*)
LOADING/UNLOADING	28					12.5	.08	1.9	
					GRAND TOTAL	155			

SYSTEM MERCURY AIR DEFENSE WEAPON SYSTEM TEST OT-II DATE 25 MAR 81 PAGE ____

NAME ____________ TELEPHONE ____________

SAMPLE

HRTES

SPI CUTOFF WORKSHEET

UNSELECTED SPI's	PRODUCTS	RELATED SYSTEM FUNCTION
PREVENTION OF INTERCEPTION/JAMMING	3.4	DESTROY AIRCRAFT
DELIVERY OF CARGO	3.1	"
ENGAGEMENT OF SEVERAL TARGETS	3.3	DENY SELECTED AIRSPACE
"	2.9	DESTROY AIRCRAFT
DELIVER AMMUNITION ON TARGET	2.9	"
"	2.1	DENY SELECTED AIRSPACE
LOADING/UNLOADING	1.9	DESTROY AIRCRAFT
MANEUVER IN TRAVEL	1.7	"
PROTECTION OF OPERATORS	1.44	"
ESTABLISHMENT & MAINTENANCE OF COMMUNICATIONS	1.4	"
MANEUVER IN TRAVEL	1.2	DENY SELECTED AIRSPACE
SELF-RECOVERY	1.2	DESTROY AIRCRAFT
INFORMATION ROUTING	1.0	"
NAVIGATION	1.0	"
MANEUVER IN TRAVEL	.80	DENY SELECTED AIRSPACE
ESCAPE FROM SYSTEM	.75	"
"	.72	DESTROY AIRCRAFT
NAVIGATION	.58	DENY SELECTED AIRSPACE
MANEUVER IN ATTACK/DEFENSE	.20	DESTROY AIRCRAFT
PREVENTION OF DETECTION/LOCATION	.12	"
LOADING/UNLOADING	.10	DENY SELECTED AIRSPACE
PREVENTION OF DETECTION/LOCATION	.08	"
ABILITY TO BE TRANSPORTED	-0-	DESTROY AIRCRAFT

SYSTEM MERCURY AIR DEFENSE WEAPON SYSTEM TEST OT-II DATE 25 MAR 81 PAGE ____

NAME ____ TELEPHONE ____

4. CONDITION CATEGORY INDEX

CONTENTS:

I. Introduction

II. Condition Category Index.

III. Combining Conditions with SPI's.

IV. Filling Out Selection Tree.

ACTIONS:

(1) Familiarization with condition categories.

(2) Rating condition categories for each SPI.

(3) Rating individual conditions, within selected categories for combination with each SPI.

(4) Development of new SPI's containing selected condition.

(5) Inclusion of SPI's and their weights in the Selection Tree.

PRODUCTS:

(1) Completed Condition Rating Worksheets.

(2) Selection Tree containing System Functions and their selection weights; SPI's and their selection weights.

USED FOR:

(1) Working papers for preparation of Independent Evaluation Plan.

(2) Working papers for Outline Test Plan.

Developing Appropriate Test Elements for This Chapter

To avoid developing test elements (System Functions, System Performance Issues, conditions, Human Performance Functions) in this chapter, when they should be developed in a later chapter, look through the following list of examples which are related. They are not complete, but they should give you an idea of the sorts of testing elements to be developed now and those which will be developed in later chapters.

SYSTEM FUNCTION EXAMPLE: (CHAPTER 2)

Destroy enemy units (aircraft, armored vehicles, fixed emplacements, personnel, etc.).

SYSTEM PERFORMANCE ISSUE EXAMPLES: (CHAPTER 3)

Ability to be transported .
Maneuver in travel.
Navigation.
Target aquisition.
Maneuver in attack/defense.
Establishment and maintenance of communications.
Delivery of ammunition on target.

YOU ARE HERE

CONDITION CATEGORY EXAMPLES: (CHAPTER 4)

Illumination.
Ground slope.
Ground surface.
Target characteristics (type, number speed, location, direction of motion, etc.
Tactical characteristics (number of friendly systems employed, speed, location, direction of motion, concealment, etc.).
Protective gear worn.
Time since end of training.
Duration of preceeding work.

HUMAN PERFORMANCE FUNCTION EXAMPLES: (CHAPTER 5)

Load system.
Unload system.
Read vehicle maneuvering instruments.
Travel designated route.
Detect target(s).
Identify target(s).
Prioritize targets.
Perform tight turn-forwards.
Establish communications net.
Load ammunition.
Aim weapon.
Fire weapon.
Reorient system to next target.

1. Introduction

At this point you will consider the conditions under which the system is to be tested. You will select conditions in two phases. First, in this chapter, you will select those conditions that are so important to a rigorous evaluation of the system that the conditions must be stated explicitly as part of a System Performance Issue. Later, in Chapter 5, you will select those additional conditions that are important for the test but that need not have such high visibility during test planning.

Up to now, the System Performance Issues (SPI's) you have selected have been general. To make them more specific you will combine them with those conditions that you will select as being critical to each SPI. HRTES supplies you with a list of conditions which are grouped into categories. These categories of conditions are listed in the "Condition Category Index" on page H4-5. A detailed list of conditions within each category is given in the "Condition Rating Worksheets," pages W4-2 through W4-41 or pages W4-47 through W4-57.

As mentioned above, the procedure for selecting conditions is done in two parts. The first part, to be completed in this chapter, consists of picking the most critical conditions from the most critical condition categories and then combining these conditions with the SPI. This process multiplies each general SPI into one or more specific SPI's. It should be noted, however, that it is not necessary to combine conditions with every SPI. Two methods are included in this chapter for combining conditions with SPI's. The first method is the more methodologically rigorous, but also more time-consuming to use.

The second part of the procedure for selecting conditions is described in Chapter 5. This second procedure focuses on those conditions that are important for the test but not so critical that they must be highlighted by combining them with SPI's. In this second procedure, the conditions are applied to the Human Performance Functions of the given SPI.

II. Condition Category Index

HRTES supplies you with a list of conditions that you will use to make the SPI's more specific to the requirements of your Operational Test. These conditions have been grouped into categories, as shown in the "Condition Category Index" on page H4-5. This index refers to selected pages in the Workbook which are the "Condition Rating Worksheets (pages W4-2 through W4-41 or pages W4-47 through W4-57. Each "Condition Rating Worksheet" includes a detailed list of conditions within the category named. The Condition Category Index is included here to give you an indication of the various categories of conditions that you will be using. As you proceed through the following instructions you will consider each category of conditions in turn and will rate each one as to the necessity of explicitly stating that category in the issues to be tested.

III. Combining Conditions with SPI's.

You will be selecting conditions for each SPI that you selected in Chapter 3. The procedure for selecting the conditions to combine with each SPI is done in two parts. The first part consists of picking the most critical categories of conditions to combine with the SPI, and then selecting at least one specific condition from each critical category. This process multiplies each general SPI into one or more specific SPI's. Note, however, that for some SPI's you may consider none of the conditions to be overwhelmingly important. It is not necessary to attach conditions to <u>every</u> SPI. The second part of selecting conditions, to be done in Chapter 5, consists of considering all other relevant categories of conditions and selecting those additional conditions which should be included in the OT.

The process of selecting conditions to combine with each SPI may be done by you with the aid of experts, or by you independently (see Introduction, page H1-9). The main part of the process consists of rating the conditions on a criticality scale. Based on the ratings obtained, you will combine the more important conditions with some SPI's to form new, specific SPI's.

CONDITION CATEGORY INDEX

Two alternative methods are provided for selecting conditions and matching them to SPI's. Both methods begin with a common General Procedure, which follows immediately. Both methods end with a second common General Procedure for determining the specific combination of conditions to apply to each SPI.

General Procedure (Initial Section)

(1) Quickly look through the condition categories and eliminate any categories that are either not within your charter to consider or that you will not be able to control. You may also want to eliminate some conditions within a single category for the same reasons. For example, it may not be within the charter of your organization to specify the range of aptitudes of the player personnel that will be used in the test. Thus, you may choose to eliminate the condition "Personnel: Aptitudes." Similarly, you may determine that you will specify certain weather conditions under which the test will be conducted, but that you will simply leave the precipitation variable uncontrolled. Thus, you would eliminate "Weather: Precipitation" from further consideration.

A convenient method for eliminating some of the condition categories is to copy the "Condition Category Index," page H4-4, and to draw a line through each condition category that you have decided to eliminate from further consideration.

At this step you are not actually "eliminating" any conditions from the test. Rather, you are deciding that these conditions will be left uncontrolled to vary "as occurs." By eliminating these conditions at this step, you will reduce the number of items that must be rated in the following steps.

(2) At this point you will rate the criticality of the condition categories, as well as the individual conditions within each category. Two general methods are presented for rating the conditions. In Method 1, condition categories and the individual conditions are rated for their relevance and criticality to each selected SPI. Therefore, each condition and condition category must be rated individually for each SPI. This method is the more rigorous of the two methods. It will probably require the use of many packages of "Condition Rating Worksheets" and a considerable amount of time to accomplish. If you have a sufficient amount of time and you do not foresee serious objections to the required workload, it is recommended that you use Method 1.

In Method 2, condition categories and the individual conditions are rated for their relevance and criticality to the tested system as a whole. Therefore, each condition and condition category need be rated only once. After the conditions have been rated for the system as a whole, you will then apply these conditions to those SPI's for which the conditions are most relevant. Method 2 is less rigorous than Method 1 since the conditions need not be considered for each SPI in turn. Thus, only one set of "Condition Rating Worksheets" need be used, rather than a complete set of Worksheets for each SPI.

Procedures for both methods are listed below. The steps for Method 1 are listed first, followed by the steps for Method 2. After you have followed the steps in either Method 1 or Method 2, then proceed to the final steps of this chapter, listed under the section entitled "General Procedure (Final Section)."

Method 1. Rating Conditions for each SPI.

(1) Copy at least one set of the "Condition Rating Worksheets" (pages W4-2 through W4-41) for each System Function, minus those condition categories eliminated as being either outside your charter, or left uncontrolled in the test. If there are more than five SPI's associated with a System Function, you will need extra sets of the "Condition Rating Worksheets."

(2) Add any condition categories that you feel are relevant, but that are not listed. Add to categories any specific conditions that you feel are relevant, but which are not listed.

(3) Enter System Function name and the statements of the selected SPI's from the SPI's worksheets in the appropriate "Selected SPI" column of the "Condition Rating Worksheet." Also fill in the name of the system and test at the bottom of each Worksheet.

(4) For each expert:

(a) Copy the prepared "Condition Rating Worksheets."

(b) Copy the "Guidelines for Selecting Conditions," pages W4-42 through W4-44, and the "Sample Worksheet," page W4-45.

Remember, *you may select yourself as an* expert. If necessary, you may be the only expert you select; however, this will greatly reduce the probable reliability.

(5) Submit the above package to the experts. Check that your package includes:

(a) Completed "Description of System to be Tested,"

(b) Completed "System Function Worksheet" (including all System Functions selected in Chapter 2),

(c) Completed "System Performance Issue Worksheet(s)" (including all SPI's selected in Chapter 3. There should be at least one Worksheet for every System Function that you selected).

(d) "Condition Rating Worksheets," and

(e) "Guidelines for Selecting Conditions" and 2-page "Sample Worksheets."

(6) When you have received the completed "Condition Rating Worksheets" from all of the experts, you will then aggregate their ratings.

For the condition category rating for each SPI, we suggest you apply the majority rule. However, the designation of a category as critical (rated "2") for an SPI will have profound influence on the field test. It is important that only the most critical categories be designated as "2". Use the "2" rating sparingly! You will not lose those conditions from categories rated "1". They can be chosen for test in Chapter 5, although they will not be attached to the SPI's. It is not necessary to attach conditions to every SPI, i.e., all condition categories for an SPI may have been rated less than "2".

For the ratings of the specific conditions within categories, we suggest that you aggregate by taking means of the experts' ratings. (Notice: the results in this case may no longer be integer numbers.) Aggregation can also be conducted by a discussion and consensus among the experts and this is the recommended method in HRTES. Record the results on fresh "Condition Rating Worksheets" (see sample Worksheet, page H4-10).

SAMPLE

HRTES

CONDITION RATING WORKSHEET

SYSTEM FUNCTION DESTROY AIRCRAFT

CONDITION CATEGORY 1. WEATHER: ILLUMINATION

SELECTED SYSTEM PERFORMANCE ISSUES

	TARGET ACQUISITION	PREVENTION OF JAMMING	ASSESSMENT OF SEVERAL TARGETS	DELIVERY OF CARGO	DELIVERY OF AMMO ON TARGET	
ILLUMINATION	2	Ø	1	Ø	2	
Full Sunlight	2		1.8		2	
Moonlight	2		1		1.5	
Starlight	0.8		1		0.5	
Dusk	0.5		1		Ø	
Overcast, Moonless Night (Pitch Black)	1.2		2		2	
Artificial Lighting (specify)	0.5		0.5		0.5	
Flares	1.1		1.2		1.5	
Direct Glare	0.5		1		Ø	
Indirect Glare (Water, Sand, Clouds, etc.)	0.5		0.5		0.5	
Other (specify)						

SYSTEM MERCURY AIR DEFENSE WEAPON SYST. TEST OT-II DATE 25 MAR 81 PAGE

NAME TELEPHONE

(7) For each SPI, consider all the condition categories that are rated "2" (i.e., critical categories). Select at least one condition from each of these critical categories. This is done based on the criticality rating of each specific condition in the category.

Note: There is a tradeoff here. Any condition from a critical category that is not selected here will not be used in the field test; however, the more conditions you choose, the more SPI's you will have.

(8) Go to the section entitled "General Procedures (Final Section)," page H4-19.

Method 2. Rating Conditions for the System as a Whole.
Remember, you may select yourself as an expert. If necessary, you may select yourself as the only expert. This would certainly be a viable alternative if you are sorely pressed for time.

If you are going to use experts to help you select the conditions, proceed to Step 3 of this section. If you have decided to select the conditions without using any other experts, you can eliminate some page-copying by proceeding directly to Step 1.

(1) Copy one set of the "Conditions x SPI Matching Worksheets," pages W4-62 through W4-70. Draw lines through any of the condition categories or individual conditions that you eliminated in Step 1 of the General Procedure, page H4-6. Add any condition categories that you feel are relevant, but that are not listed.

You will now use these "Conditions x SPI Matching Worksheets" for two related purposes. First you will rate the conditions for the system as a whole and record these ratings in the first column of the worksheets. Next, you will match these ratings for the conditions with specific SPI's.

(2) Use the instructions in the "Guidelines for Selecting Conditions - Method 2" to select the conditions for the system as a whole. Just remember that you are recording your selections in the first column of the "Conditions x SPI Matching Worksheet," rather than on the "Condition Rating Worksheet - Method 2." (This eliminates some recopying that would otherwise need to be done.)

Go to Step 9 of this section, page H4-15.

(3) Copy one set of the "Condition Rating Worksheets Method Two," pages W4-47 through W4-57. Draw lines through condition categories or individual conditions that are either outside your organization's charter, or to be left uncontrolled in the test.

(4) Add any condition categories that you feel are relevant, but that are not listed. Add to categories any specific conditions that you feel are relevant, but that are not listed.

(5) Fill in the name of the system and test at the bottom of each Worksheet.

(6) For each expert:

(a) Copy the prepared "Condition Rating Worksheet - Method 2."
(b) Copy the "Guidelines for Selecting Conditions - Method 2," pages W4-58 through W4-61.

(7) Submit the above package to the experts. Check that your package includes:

(a) Completed "Description of System to be Tested,"
(b) Completed "System Function Worksheet" (including all System Functions selected in Chapter 2),
(c) Completed "System Performance Issue Worksheet(s)" (including all SPI's selected in Chapter 3. There should be at least one Worksheet for each System Function that you selected).
(d) "Condition Rating Worksheets - Method 2," and
(e) "Guidelines for Selecting Conditions - Method 2" and the 2-page "Sample Worksheets - Method 2."

(8) When you have received the completed "Condition Rating Worksheets - Method 2" from all of the experts, you will then aggregate their ratings.

You should be aware that the designation of a condition category as critical (rated "2") implies that one or more of the conditions in that category will definitely be selected for combination with one or more of the SPI's. Such a combination will have a profound influence on the field test. Only the most critical condition categories should be designated as "2". Those condition categories rated as "1" will not be combined with SPI's. However, you will consider the "1"-rated condition categories later in Chapter 5.

Three methods for aggregating these ratings of conditions are suggested:

(a) The recommended method is aggregation through direct discussion and consensus among those experts who made the ratings. This method is best applied immediately following the rating process and without any time break. It takes the form of the various experts discussing their original decisions, as necessary, and arriving at single rating for each category and condition.

(b) The next possible method for aggregation is the majority rule. In this method, the majority of ratings ("0", "1", or "2") determines the rating of each condition and category.

(c) The third possible method of aggregation is the computation of means for each category and condition, and the use of the resulting means as a decision aid for you in assigning "0", "1", or "2" ratings to each condition and category. In this method, you would compute means for each rating. You would then have to decide, based on each resulting mean, and the relationship between them, which should be assigned "0", "1", or "2" ratings.

Remember, this whole rating process is designed as a decision aid for you to select the items that will be important to include in the test plan. Since it is ultimately your decision, you can select the method that gives you the most aid.

(9) Once you have decided which condition categories and which individual conditions should be rated "0", "1", or "2", you now have to match the conditions with the appropriate SPI's. Remember, these conditions and condition categories have been selected for the system as a whole. They must now be applied to those SPI's selected in Chapter 3. To aid you, a "Conditions x SPI's Matching Worksheet" has been included on pages W4-62 through W4-70. Use the following steps to complete this worksheet.

(10) Copy the "Conditions x SPI Matching Worksheet," pages W4-62 through W4-70. There are spaces for matching all listed conditions with 20 SPI's on the Worksheet. If you have selected more than 20 SPI's you will have to make more than one copy of each page of the Worksheet.

(11) Retrieve your copies of the "System Performance Issue Worksheet(s) that you completed in Chapter 3. You have marked all SPI's to be included in the test plan with asterisks on these Worksheets.

(12) Assign identification numbers to each SPI marked with an asterisk and to each System Function from which the SPI's are derived. (This is done so that the resulting System Function and SPI identification numbers can be listed in the restricted spaces available on the "Conditions x SPI Matching Worksheet.")

(13) List the SPI ID number of each selected SPI, and the parent System Function's ID number, in the two upper rows of the "Conditions x SPI Matching Worksheet."

(14) If an entire condition category has not be rated at all, because it was outside your organization's charter, or it is to be left uncontrolled in the test, put a line through it on the Worksheet.

(15) Based on your aggregation of the experts' ratings, each remaining condition category has been given a "0", "1", or "2" rating for the system as a whole. If a condition category has been given an aggregated rating of "0", all SPI's will receive a "0" for that condition category, and all the conditions in that category may be ignored. If a condition category has been given an aggregated rating of "2", some or all SPI's will receive a "2" for that condition category. Remember, when you match a condition category rated "2" to a specific SPI, one or more of the conditions in that category will be linked to one or more SPI's to produce one or more new SPI's which replace the original ones. Therefore, the decisions to match a given SPI with a category rated "2" must be made quite carefully. The following decision order is suggested:

(a) For condition categories rated "0", record 0's in all SPI cells. Individual conditions within such categories will not be considered further, but will be left uncontrolled to vary "as occurs" in the test.

(b) For condition categories rated "2", record either "2", "1", or "0" in each SPI cell. At least one SPI must receive a "2", and as many SPI's may receive a "2" as you think appropriate. However, you should remember that recording a "2" for an SPI means eventual combination of that SPI with one or more conditions in that category.

Recording a "1" means that in Chapter 5, those conditions will be considered for the Human Performance Functions derived from that SPI. Recording a "0" means the conditions in that category will not be considered either for that SPI, or for the Human Performance Functions derived from the SPI.

(c) For condition categories rated "1", record either "1", or "0" in each SPI cell. At least one SPI must receive a "1", and as many SPI's may receive a "1" as you think appropriate. No SPI's may receive "2's" for this condition category, but they may receive "0's."

You have now assigned a "0," "1," or "2" to each condition category cell for each SPI. You will now assign ratings to each individual condition within each category. If a "0" has been assigned to a condition category for a given SPI, it means that you will allow that condition category to vary "as occurs" for that SPI. Thus, all individual conditions within that category are also assigned a "0."

If you have assigned a "2" to a condition category cell for a given SPI, it means that at least one individual condition will be combined with that SPI. Thus, at least one individual condition within the category must be assigned a "2." It is suggested that you use the experts' ratings of the individual conditions for the system as a whole in assigning the ratings of the individual conditions for the given SPI. For example, if the condition category "Weather: Illumination" had been rated "2" for the SPI "How effectively can the system maneuver on the battlefield?", you would use the experts' ratings to rate the individual conditions, as shown:

Example:

CONDITIONS	AGGREGATED EXPERTS' RATINGS (system as a whole)
Full Sunlight	0
Moonlight	0
Starlight	2
Dusk	0
Pitch Black	0
Artificial Lighting	0
Flares	0
Direct Glare	0
Indirect Glare	0
As Occurs	2

Thus, the individual conditions "Starlight" and "As Occurs" would be rated "2" for this SPI. The result is two new SPI's as follows:

1. "How effectively can the system maneuver on the battlefield under starlight conditions?"
2. "How effectively can the system maneuver on the battlefield with illumination as it occurs during the test?"

Notice: By assigning a "2" to the condition "As Occurs," you have stated explicitly that illumination will be tested "as occurs", in addition to testing under "Starlight" conditions.

If you have assigned a "1" to a condition category cell for a given SPI, it means that at least one of the individual conditions will be considered later in Chapter 5 as part of a Human Performance Function for that SPI. Thus, at least one individual condition within the condition category must be assigned a "1." No individual condition may be rated "2," but some individual conditions may be rated "0."

General Procedures (Final Section)

(1) For each SPI, list the conditions which will be attached to it.

The conditions should be divided into their categories. For example, suppose that for the SPI "How effectively can the system acquire its target?" two condition categories (Illumination and Target Location) were regarded as critical. For each one, two specific conditions were selected. The list would be written as follows:

SYSTEM FUNCTION: DESTROY AIRCRAFT

SPI: TARGET ACQUISITION

Specific Condition	Condition Category	
	ILLUMINATION	TARGET LOCATION
1	Full Sunlight	Maximum range
2	Moonlight	Normal Range

(2) Determine the combination of conditions selected and attach them to the SPI.

There are many possible combinations of conditions. Choosing one condition from each category rated as "2" yields one combination. For example, using the matrix above, there are four possible combinations as follows:

(a) Full sunlight and maximum range.
(b) Moonlight and maximum range.
(c) Full sunlight and normal range.
(d) Moonlight and normal range.

A new SPI would be made by combining the general SPI with a selected combination, such as "How effectively can the system acquire the target in full sunlight and maximum range?"

If all four combinations are so important as to be required for the OT, the general SPI would then multiply into four new SPI's. You may decide, however, that not all of the combinations are sufficiently important to be included in the OT. You may choose two or three of the combinations. Note that you must choose at least two combinations to assure that all four specific conditions are represented. Thus, combinations (a) and (d), or combinations (b) and (c) would be selected. Note that any combination not selected will not be included in the OT.

The combination of an SPI and its condition(s) will now be referred to as an SPI. Those SPI's to which conditions are not attached will also be referred to as SPI's.

It is suggested that the decisions required for combining SPI's and conditions be made in concert with the experts who selected the conditions, if possible.

(3) You have now developed all the SPI's for the OT. You should record them on the "SPI Summary Worksheet" for further use. The following is a guidance for filling in the "SPI Summary Worksheet" (see "Sample SPI Summary Worksheet," page H4-22).

(a) For each System Function make one copy of the "SPI Summary Worksheet" found on page W4-71.

(b) Write in the System Function name.

(c) Write in the first column the SPI statements (without the conditions attached to them).

(d) Record the selection weights on the second column. These weights are found on the completed "System Performance Issue Worksheet."

(e) Record the condition combination attached to each of the SPI's, in the third column.

(f) Column 4 will be filled in following section of this chapter.

SAMPLE

HRTES

SPI SUMMARY WORKSHEET

SYSTEM FUNCTION: DESTROY AIRCRAFT

SPI (STATEMENT)	SELECTION WEIGHT	CONDITION COMBINATIONS ATTACHED TO SPI STATEMENTS	FINAL SELECTION WEIGHTS
TARGET ACQUISITION	23.5	FULL SUNLIGHT, MAXIMUM RANGE (5 MILES)	11.75
		MOONLIGHT, NORMAL RANGE (2 MILES	11.75
PREVENTION OF INTERCEPTION/JAMMING	21.5	NONE	21.50
DELIVERY OF AMMO ON TARGET	19.0	FULL SUNLIGHT	9.50
		OVERCAST MOONLESS NIGHT	9.50
... ETC.			

SYSTEM MERCURY AIR DEFENSE WEAPON SYSTEM TEST OT-II DATE 25 MAR 81 PAGE ____

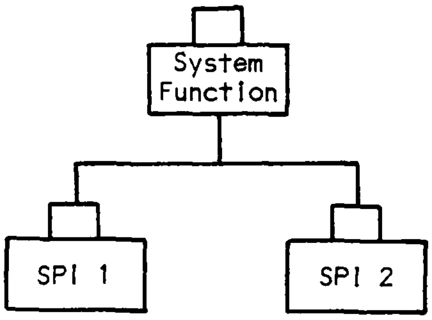
NAME ____________________ TELEPHONE ____________________

IV. Filling Out the Selection Tree.

You should now expand your Selection Tree which you started in Chapter 2 to include one more level: The SPI Level. Each System Function on the Tree will branch into a number of nodes, one for each SPI.

(1) On your "Selection Tree" draw nodes for the SPI's to be included.

For each System Function draw a node for each SPI which belongs to it. This information is recorded on the "SPI Summary Worksheets" you developed in the previous section. A System Function with just two SPI's will be structured as follows:



(2) Write in the appropriate space of each node the name of the SPI (remember the names of some SPI's include the conditions selected in this chapter).

(3) Determine and record the final selection weight of each SPI.

The original Selection Weights are recorded in the "SPI Summary Worksheets." Those SPI's which were not combined with conditions maintain their original selection weights. SPI's which were created by attaching conditions to original SPI's must have final selection weights computed for them. To do this, divide the original selection weight of the SPI by the number of new SPI's produced from it. For example, if an original SPI had a weight of 30, and if it has been combined with conditions to form two SPI's, each of the new SPI's will have a final selection weight of 30 ÷ 2 or 15. Record the final selection weights in the fourth column of the "SPI Summary Worksheets." Also record the final selection weights on the Selection Tree.

A Technical Suggestion for Drawing the Selection Tree:

If the Selection Tree becomes so large it is not possible to contain it all, even on a reasonably large sheet of paper, divide the Tree in two. The first levels of the Tree which includes: System, System Functions, and SPI's would be drawn on one piece of paper. HPF-Groups and the levels underneath them would be drawn on separate sheets of papers for each SPI. Thus, each SPI and the nodes which branch from it would be described on a spearate piece of paper. Following this method, the number of pieces of paper needed to contain one Selection Tree would be equal to the number of SPI's + 1.

NOTE: Since you have a group of knowlegeable experts who developed the "Criterion Worksheets" it is recommended that you now perform the procedures in Chapter 7 section III which develops a Value Function for each statistic. The procedure in these sections should be performed by experts with a knowledge of the HRTES performance and statistic criteria.

5. HUMAN PERFORMANCE FUNCTION GROUPS (HPF-Groups)

CONTENTS:

I. Selecting Operational Human Performance Function Groups (HPF-Group) and their individual HPF's.

II. Developing Maintenance Human Performance Functions.

III. Attaching Conditions to Human Performance Functions.

IV. Developing Performance Criteria.

V. Developing Statistics and Statistic Criteria.

VI. Filling Out Selection Tree.

ACTIONS:

(1) Selecting Operational Human Performance Functions.

(2) Developing Maintenance Human Performance Functions.

(3) Developing sets of conditions for testing each HPF (either maintenance, or operation).

(4) Developing performance measure(s) and performance criterion for each HPF and its test conditions.

(5) Developing statistic(s) and statistic criterion for each measure.

(6) Inclusion of new HPF-Groups, HPF's and statistics on the Selection Tree.

(7) Computing weight for each node developed in this chapter and recording it on the Selection Tree.

PRODUCTS:

(1) Completed Operational HPF-Group Worksheets.

(2) Completed Maintenance Worksheets.

(3) Completed Operational Test Condition Worksheets.

(4) Completed Final Condition Set Worksheets.

(5) Completed Criterion Worksheets.

(6) Completed Selection Tree.

USED FOR: (1) Working papers for the Independent Evaluation Plan.

(2) Working papers for the Outline Test Plan.

(3) Working papers for the Test Design Plan.

(4) Working papers for the Detailed Test Plan.

Developing Appropriate Test Elements for This Chapter

To avoid developing test elements (System Functions, System Performance Issues, conditions, Human Performance Functions) in this chapter, when they should be developed in a later chapter, look through the following list of examples which are related. They are not complete, but they should give you an idea of the sorts of testing elements to be developed now and those which will be developed in later chapters.

SYSTEM FUNCTION EXAMPLE: (CHAPTER 2)

Destroy enemy units (aircraft, armored vehicles, fixed emplacements, personnel, etc.).

SYSTEM PERFORMANCE ISSUE EXAMPLES: (CHAPTER 3)

Ability to be transported.
Maneuver in travel.
Navigation.
Target aquisition.
Maneuver in attack/defense.
Establishment and maintenance of communications.
Delivery of ammunition on target.

CONDITION CATEGORY EXAMPLES: (CHAPTER 4)

Illumination.
Ground slope.
Ground surface.
Target characteristics (type, number speed, location, direction of motion, etc.
Tactical characteristics (number of friendly systems employed, speed, location, direction of motion, concealment, etc.).
Protective gear worn.
Time since end of training.
Duration of preceeding work.

YOU ARE HERE

HUMAN PERFORMANCE FUNCTION EXAMPLES: (CHAPTER 5)

Load system.
Unload system.
Read vehicle maneuvering instruments.
Travel designated routes.
Detect target(s).
Identify target(s).
Prioritize targets.
Perform tight turn-forwards.
Establish communications net.
Load ammunition.
Aim weapon.
Fire weapon.
Reorient system to next target.

I. Selecting Operational Human Performance Function Groups (HPF-Groups) and the Individual HPF's.

Human Performance Functions (HPF's) are those actions performed by humans, with or without system hardware, which are required to perform a given SPI. In most cases, there are two general classes of HPF's required for the performance of an SPI. Those HPF's required for the actual performance of the SPI are called Operational HPF's. Those required for continued performance of an SPI are called Maintenance HPF's. HPF's are collected in HPF-Groups to aid you in considering the individual HPF's, to simplify the process of developing criteria and measures, and to structure the evaluation process.

This section will address the selection of Operational HPF's. Section II of this chapter will address the development of Maintenance HPF's.

The following instructions will guide you in selecting specific Operational HPF's for testing:

(1) Review your "System Performance Issue Worksheets" (Chapter 3), and find the HPF-Group numbers referenced by each selected SPI.

(2) Use the referenced HPF-Group number to locate the corresponding "Operational HPF-Group Worksheets," which begin on page W5-2. The HPF-Group numbers are located in the upper left of each Worksheet.

(3) If more than one Operational HPF-Group has been referenced by a given SPI, determine if some of the referenced Groups are not applicable to your system. It is permissible to select more than one Operational HPF-Group per SPI, if you so decide.

(4) Copy the applicable "Operational HPF-Group Worksheet(s)," and add any other Human Performance Functions (HPF's) that you feel are relevant to your HPF-Group but which are not listed. (See Sample Worksheet, page H5-6).

(5) If you think a significant Operational HPF-Group is missing for a given SPI, you will have to prepare it using existing task analyses, experts, or your own expertise. Be sure to use the HRTES format in stating the HPF's. Insert the new HPF-Group in the Workbook, and apply all other HRTES procedures to it exactly as if it had appeared in HRTES originally.

(6) For each HPF-Group, select those HPF's that are relevant to the performance of the SPI and mark them in the appropriate column of the Worksheet. By relevant we mean:

   (a) The HPF is expected to be performed as part of the SPI.
   (b) The level of performance of the HPF is expected to have an effect on the level of system performance.

(7) Select HPF's for testing.

Potentially, each relevant individual HPF will be tested. However, if you think it is more appropriate to consider two or more HPF's (in the same HPF-Group) as one, you can combine them together into a single HPF. The last column of each "Operational HPF-Group Worksheet" is provided so that you can assign new numbers to the HPF's selected. Assign the same number to HPF's you wish to combine.

SAMPLE

HRTES

OPERATIONAL HPF-GROUP WORKSHEET

SYSTEM FUNCTION DESTROY AIRCRAFT

SPI TARGET ACQUISITION

HPF-GROUP # 2 TARGET ACQUISITION

HPFs	MARK IF RELEVANT	NEW HPF NUMBER
1. Detect target(s)	X	1
2. Identify target(s)	X	1
3. Select target(s) and target order	X	2
4. Orient weapon system in general firing position		
5. Determine range of target	X	3
6. Aim weapon system. This involves a procedure which results in the system being adjusted for the azimuth and elevation of the target		
7. Illuminate or designate target		
8. Adjust aim, following miss		
9. Shift to second target	X	5
10. MATE MISSILE WITH APPROPRIATE WARHEAD: NUCLEAR/CONVENTIONAL	X	4

SYSTEM MERCURY AIR DEFENSE WEAPON SYSTEM TEST OT-II DATE 25 MAR 81 PAGE ____

NAME __________ TELEPHONE __________

You might combine HPF's if:

(1) It is more convenient to measure performance on the combined HPF than on each HPF separately, or

(2) You find it more meaningful to consider several small HPF's together.

In some cases, you may even consider combining all HPF's in an HPF-Group into one single HPF for measurement purposes. HPF's combined together are now considered as one HPF, equivalent to a single HPF which was not combined with any other one.

II. Developing Maintenance Human Performance Functions.

When a system breaks down in OT, time and accuracy measures of maintenance activities, plus questioning of the participants, should take place. Often, however, the OT will be relatively short and will not stress the system in a completely realistic way. For this reason, the kinds of system breakdowns which occur in battlefield use may not take place in the OT, but they should be simulated to fully evaluate the system. Therefore, it is important to include relevant Maintenance HPF's in the Operational Test.

It is not possible for HRTES to supply a comprehensive list of all Maintenance HPF's for all types of systems. Therefore, HRTES includes procedures by which you can obtain Maintenance HPF's from maintenance experts. A form and instructions are provided which you are to send to individual maintenance experts. The experts must be knowledgeable about the specific maintenance requirements and anticipated problems of the system, or similiar systems. It is recommended that more than one expert be involved with developing Maintenance HPF's for each SPI. It is not necessary, however, for each expert to be involved with all SPI's.

(1) For each expert, make:

(a) One copy of the "Guidelines for Developing Maintenance Human Performance Functions" (pages W5-46 and W5-47).

(b) Sufficient copies of "Maintenance Worksheets" (page W5-48) to cover all selected SPI's. More than one Worksheet may be required for each SPI.

(2) Fill in the required information on the top of the "Maintenance Worksheets." If more than one expert is to be used for a given SPI, copy one set of Worksheets for each expert.

(3) Submit Guidelines and Worksheets to the appropriate experts.

(4) After receiving the completed Worksheets, produce a final set of Maintenance HPF's for each SPI.

If only one maintenance expert provided input to this process, the resulting completed Worksheets will serve as the basis of the Maintenance HPF-Groups. However, if more than one maintenance expert is involved in this process, it will be necessary to combine the resulting data for each SPI. This may be done by agreement of experts based on the Worksheets obtained, or by combining essentially similar HPF's and selecting those which appear most frequently. If the latter alternative is chosen, it is recommended that you consult with experts regarding the accuracy and completeness of the final report.

At this point you, in consultation with experts, can combine individual Maintenance HPF's as you did for Operational HPF's. Refer to Section I for guidance. All the Maintenance HPF's which are required for a given SPI are to be considered a Maintenance HPF-Group.

(5) Record the final list of Maintenance HPF's on a fresh "Maintenance Worksheet(s)."

III. Attaching Conditions to Human Performance Functions.

In Chapter 4, you selected critical conditions which were combined with SPI's to form new, more specific SPI's. Now it is necessary to specify complete sets of conditions for the Operational Test. You must specify the conditions under which each HPF will be tested. Several HPF's will probably be tested under the same set of conditions.

You will first be asked to collect all those HPF's (Maintenance and Operational) within an SPI for a given System Function which will be tested under the same set of conditions. For each of these, you will first be asked to specify the conditions that were already selected for its SPI, to select additional conditions, and then to specify sets of these conditions for the test.

It is probable that all the HPF's in a given *Operational HPF-Group will* be performed in the same set(s) of conditions. However, HPF's in Maintenance HPF-Groups are more likely to be performed in differing sets of conditions. That is, Maintenance HPF's performed by maintenance personnel in maintenance facilities are likely to be performed in different condition than those performed by system operators in the field, even if both types of Maintenance HPF are present in the same HPF-Group.

In Chapter 4, conditions were organized into categories. Some of the categories were considered irrelevant to a specific SPI (i.e., rated "0"). Other categories were considered so critical (i.e., rated "2") that one or more of their conditions were attached to the SPI. At this point you 'll reconsider those categories of conditions which were rated "1" and select at least one specific condition from each category under which the HPF will be tested during the Operational Test. Note that those

condition categories previously rated "2" are now part of the SPI and will automatically be included in the OT.

You can do the following steps alone or in consultation with experts. HRTES supplies you with instructions and forms for obtaining the necessary information from the experts.

For each SPI within each System Function, do the following:

(1) Copy a sufficient number of "Test Condition Worksheets" (page W5-49). You will need at least one copy for each SPI within each System Function. Note, there is space for only five HPF's per Worksheet. Each sheet should contain only those HPF's which will be tested under the same conditions.

(2) Decide which HPF's (Operational and Maintenance) will be tested under the same set of conditions (whatever those conditions will be).

Usually HPF's which are performed together will be tested under the same conditions. In many cases, these will be the HPF's in an HPF-Group. It is likely that some Maintenance HPF's will not be tested under the Condition(s) specified in their parent SPI.

For example: If the parent SPI were -- "How effectively does the system acquire targets in a sandstorm?" -- it is unlikely that scheduled maintenance by maintenance personnel would take place in a sandstorm. However, unscheduled maintenance by system operators is quite likely to take place in the sandstorm. Therefore, the unscheduled Maintenance HPF would be placed on the same "Test Condition Worksheet" as were the Operational HPF's. The scheduled Maintenance HPF would appear on a separate "Test Condition Worksheet."

(3) Fill in the top of each Worksheet, and list the HPF's which will be tested under the same conditions in the appropriate spaces on each Worksheet. The HPF's are found in the "Operational HPF-Group Worksheets" and "Maintenance Worksheets" you prepared in this chapter.

(4) List the relevant condition categories for each SPI within each System Function. This information is contained in the "Condition Rating Worksheets" you completed in Chapter 4.

(a) Find the "Condition Rating Worksheets" which correspond to the SPI and System Function you recorded at the top of each "Test Condition Worksheet."

(b) List the condition categories rated "1" or "2" from the appropriate "Condition Rating Worksheets" in the first column of the "Test Condition Worksheet."

(5) List the condition(s) already included in the SPI in the appropriate places of the Worksheet. (See the sample "Test Condition Worksheet;" page H5-12).

In the case of those Maintenance HPF's which you have decided will not be performed under the conditions included in their SPI, disregard this instruction.

(6) Assign one or more specific condition from each remaining condition category and write it in the appropriate spaces of the Worksheets.

The specific conditions chosen should be based on the ratings you recorded on your "Condition Rating Worksheet" in Chapter 4.

SAMPLE

HRTES

TEST CONDITION WORKSHEET

SYSTEM FUNCTION	DESTROY AIRCRAFT
SPI	TARGET ACQUISITION IN FULL SUNLIGHT
HPF's PERFORMED UNDER THE SAME CONDITIONS	DETECT AND IDENTIFY TARGETS
	SELECT TARGETS AND TARGET ORDER
	AIM WEAPON SYSTEM

RELEVANT CONDITION CATEGORIES	SELECTED CONDITION(S) FROM EACH CATEGORY
ILLUMINATION	1) FULL SUNLIGHT
TEMPERATURE	1) HIGH (> 90°F)
	2) NORMAL (50°-90°F)
PRECIPITATION	1) NO PRECIPITATION
TARGET TYPE	MIG 25
TARGET NUMBER	1) SINGLE
	2) MULTIPLE (10 TARGETS + 15 NON-TARGET A.C.)
TARGET LOCATION	1) NORMAL RANGE (2 MILES)
	2) MAXIMUM RANGE (5 MILES)
TARGET SPEED	600 - 900 MPH.
LENGTH OF PRECEEDING WORK	1) FOLLOWING EXTENDED PERIOD OF WORK (20 HRS.)
	2) FOLLOWING NORMAL PERIOD OF WORK (3-5 HRS.)
ETC.	

SYSTEM MERCURY AIR DEFENSE WEAPON SYST. TEST OT II DATE 25 MAR 81 PAGE ____

NAME ____________________ TELEPHONE ____________

(7) Examine each completed "Test Condition Worksheet," and decide which set(s) of conditions to apply to all the listed HPF's. This may be done by you independently or through agreement with the appropriate experts.

You have now selected all the conditions under which you want to test each HPF. You now must decide in what way to combine these conditions, in other words, to determine the condition set(s) for each HPF.

Each condition set for an HPF must include one and only one condition from each relevant condition category. Having two conditions from the same category forces you to have at least 2 condition sets. However, you do not have to choose every possible combination of conditions, and in fact, it is highly unlikely that you will do so. A condition set should not be inconsistent or so unique that it is unlikely that the system would ever encounter it. For example: Rain (precipitation) and low humidity (Humidity) are inconsistent. The condition sets you select for an HPF should be those which are significant for its evaluation. It is suggested that you consider including condition sets which represent the normal operational situation, worst probable operational situation, and best probable operational situation.

(8) Copy a sufficient number of "Final Condition Set Worksheets," (page W5-50) so that you have space to record each HPF and its condition set(s). There will be at least one Worksheet for each SPI of each System Function.

(9) Fill in the top of a Worksheet and list each HPF and the condition set(s) under which it will be performed. For clarity, it is desirable to keep the same HPF's together on a "Final Condition Set Worksheet" which were listed together on a "Test Condition Worksheet." (see Sample "Final Condition Set Worksheet, page H5-14).

(10) Assign each condition set a number and record it in the first column of the "Final Condition Set Worksheet."

IV. Developing Performance Criteria.

You have now reached the level of HPF's in the decomposition of the system's performance. At this level, measurement and evaluation of actual human performance will take place. At this level one would like to determine when an HPF trial is performed successfully. To do so, it is necessary to define what one successful HPF trial means. In other words, what is the performance criterion for each HPF. Usually this can be done by determing the maximum acceptable time to perform the HPF, and/or by specifying some minimum level of performance accuracy. The procedure which HRTES suggests is, first to develop a performance criterion for each HPF, and then from each criterion to derive the appropriate measure to be taken.

There are three types of criteria to be considered: (1) time criterion, (2) accuracy criterion and (3) combined time and accuracy criterion. It is expected that the third type will be used most frequently. Examples of these criteria are: Time to perform an HPF should not exceed 30 seconds; Number of errors while performing this HPF should not exceed 10; HPF should be done within 10 seconds and with no more than 5 errors.

The measures that follow from those performance criteria are: (a) Time to perform the HPF, (b) Number of defined errors in the HPF performance, and (c) Time to perform the HPF and number of defined errors in the performance.

SAMPLE

HRTES

FINAL CONDITION SET WORKSHEET

SYSTEM FUNCTION DESTROY AIRCRAFT

SPI TARGET ACQUISITION IN FULL SUNLIGHT

HPF's PERFORMED TOGETHER
- DETECT & IDENTIFY TARGETS
- SELECT TARGETS & TARGET ORDER
- AIM WEAPON SYSTEM

SET #	CONDITION SET(S) UNDER WHICH EACH HPF IS TO BE TESTED
1	FULL SUNLIGHT, NORMAL TEMPERATURE, NO PRECIPITATION,
	SINGLE MIG 25, AT 2 MILES, 600-900 MPH,
	FOLLOWING NORMAL PERIOD OF WORK (3-5 HRS), ETC.
2	FULL SUNLIGHT, HIGH TEMPERATURE, NO PRECIPITATION,
	10 SIMULTANE 'S MIG TARGETS PRESENTED WITH 15
	OTHER AIRCRAFT, NORMAL RANGE, 600-900 MPH,
	FOLLOWING 20 HRS OF WORK, ETC.
3	SAME AS 1 EXCEPT SUBSTITUTE MAXIMUM RANGE
	FOR NORMAL RANGE.
	ETC.

SYSTEM MERCURY AIR DEFENSE WEAPON SYST. TEST OT-II DATE 25 MAR 81 PAGE ____

NAME ____________ TELEPHONE ____________

Frequently, when specifying an accuracy criterion, it is useful to distinguish between different kinds of errors which can occur during the HPF. In this case, you will have to specify the maximum number of errors allowed for each error type.

While developing a performance criterion for an HPF, the condition set(s) under which the HPF will be tested should be taken into consideration. It is expected that the measures for an HPF will be the same under various condition sets. However, under different condition sets, different performance criteria may be appropriate.

The following procedure may be done by a group of experts including you, or by you independently. Due to the importance of receiving acceptance for the product of this procedure, it is recommended that criteria be developed with experts. Experts selected should be knowledgable about required tactical constraints on system operation and maintenance. All the information which will be developed in this section will be recorded on the "Criterion Worksheet."

HRTES supplies you with forms and instructions to be sent to experts. The following instructions will aid you in developing appropriate criteria.

(1) For each expert:

  (a) Copy the "Guidelines for Developing Performance Criteria" (page W5-51).
  (b) Copy the "Final Condition Set Worksheets" for the appropriate HPF's that you completed in the last section.
  (c) Make a sufficient number of copies of the "Criterion Worksheets" (page W5-57) and attach them to the "Final Condition Set Worksheets."

(a) Record on each "Criterion Worksheet" the condition set numbers from the "Final Condition Set Worksheet."

(b) Copy the "Sample Criterion Worksheet" (pages W5-55and W5-56).

(2) Submit the above items to each expert.

If you are identifying the performance criteria without the assistance of experts, use one set of materials that would have been submitted to the experts and perform the steps described in the "Guidelines for Developing Performance Criteria."

(3) After you have received the completed Worksheets from the experts for all the HPF's, you will have to aggregate the resulting data from the various experts onto fresh Worksheets for each HPF. This may be done either with the consensus of the experts, using the Worksheets as the basis, or by (a) consolidating similar error, (b) computing means of each time criterion, (c) selecting only those error that a majority of experts included in their criterion for the HPF, in its specific condition set.

(4) After the process of aggregation is completed, record the results on fresh "Criterion Worksheets."

(5) You have now defined successful performance of a single trial in terms of time and accuracy. The criterion of a successful trial must, however, also include the following statement:

> "If a <u>significant</u> accident or near accident occurs involving personnel or hardware as a direct result of the performance of a trial of this HPF, that trial is to be judged below criterion."

V. Developing Statistics and Statistics Criteria.

At this point a definition has been made for one successful trial of each HPF under each condition set. These are the HPF performance criteria. Since the evaluation process will not be reliable if based on one trial for each HPF under each condition set, several trials for each HPF will have to be performed. Chapter 6 gives a detailed description of the process for determining the number of trials of each HPF to perform.

The concern of this section is how to aggregate the data obtained from many HPF trials, and whether or not to aggregate data from an HPF separately under its various conditions sets.

There are two primary ways of aggregating the data: (a) by taking the average of all the measure outcomes for the HPF, (b) by calculating the percentage of successful trials of the HPF. One successful trial is defined by the performance criterion of the HPF (see Section IV). HRTES refers to the average, percentage, or any other possible aggregation methods as a statistic. The average and the percentage can be expressed by the following formulas:

$$\text{AVERAGE} = \frac{\text{Sum of total outcomes}}{\text{Number of trials}}$$

$$\text{PERCENTAGE} = \frac{\text{Number of successful trials}}{\text{Number of trials}} \times 100$$

If you are interested in including condition sets to insure that the HPF is performed under representative conditions rather than to determine specific effects of condition sets on HPF performance, aggregation should be done across condition sets. However, if you are principally interested in the effects of each condition set on HPF performance, aggregation should be done for each condition set of the HPF. In the latter case considerably more trials of the HPF will be required to insure reliability.

In this section, you are asked to determine which kind of statistic to use for aggregation of the data from the various trials of each HPF, and whether aggregation should be across condition sets or done separately for each condition set. The latter decision will considerably affect the number of trials needed for each HPF under each condition set (a detailed discussion of this is given in Chapter 6).

To evaluate the performance of HPF's, you need to define a criterion for each statistic that is employed. This criterion is called the statistic criterion. For example, a statistic criterion for a percentage statistic might be -- "A minimum success percentage of 80%." A statistic criterion for an average statistic might be: "At most, an average time of 30 seconds."

All the information which is developed in this section will be recorded on the "Criterion Worksheet."

The following procedure may be done by a group of experts including you, or by you independently. It is highly recommended that in the former case the experts to be selected be the same experts you used for determining the performance criteria, since they are familiar with the processes and understand these criteria.

(1) For each expert:

(a) Copy the "Criterion Worksheets" which were prepared in step 4 of the previous section.

(b) Copy one set of appropriate "Final Condition Set Worksheets" (the same set that was copied in step 1b of the previous section).

(c) Copy the "Guidelines for Developing Statistics and Statistic Criterion." (Pages W5-58 to W5-61)

(d) Copy the "Sample Criterion Worksheet" (pages W5-62 and W5-63).

(2) Submit the above items to each expert.

If you are developing the statistics and statistic criteria without assistance of experts, use one set of the above items and perform the steps described in the "Guidelines for Developing Statistics and Statistic Criterion."

(3) After you have received the completed Worksheets from the experts for all HPF's, you will still have to aggregate all the information from various experts and to record the result on fresh Worksheets.

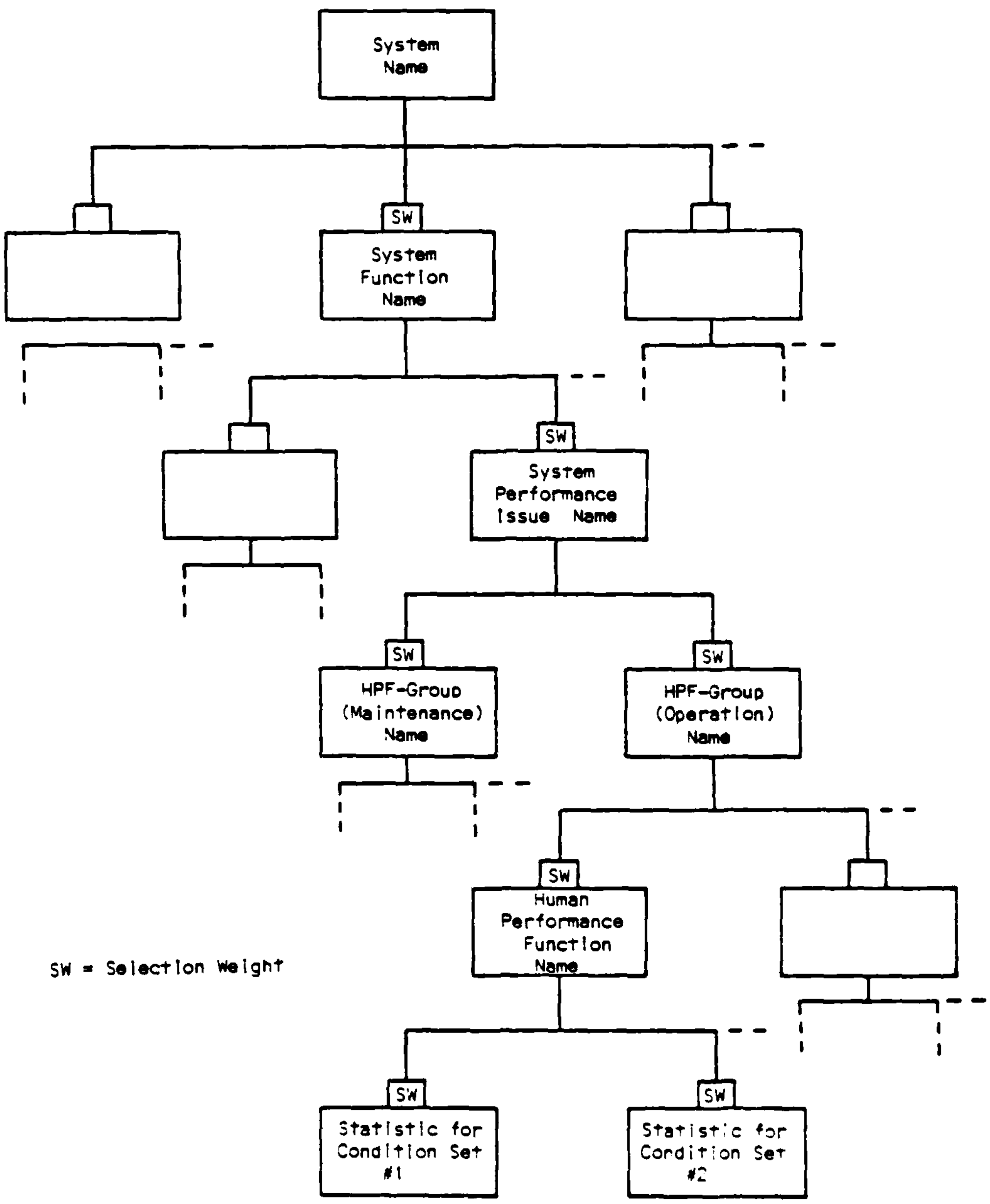
The aggregation may be done in the usual manner either by taking means, majority rule or by consensus of the experts. If not all the experts have agreed whether an HPF should have one statistic calculated across all condition sets or calculated for each of its condition sets, you will have to make this decision.

(4) Now you should have a completed "Criterion Worksheet" for each HPF and thus a criterion, a description of time and errors to be taken, a statistic-criterion, and a statistic for every HPF under each condition set.

VI. Filling Out the Selection Tree.

At this point you should put all of the information obtained in this chapter on the Selection Tree. So far the Selection Tree consists of three levels: level 1: System, level 2: System Function, and level 3: SPI's. Each of the SPI's (with or without conditions included) will be further decomposed to HPF-Groups. Each HPF-Group will be decomposed to HPF's and the HPF's wil be decomposed to the appropriate statistics. The bottom of the tree will consist of the statistics. Schematically, it will look as shown in Figure H5-1 (page H5-21).

Figure H5-1
SELECTION TREE STRUCTURE



The required information for you to fill out the Selection Tree is already recorded on the following Worksheets that you completed in this chapter:

"Maintenance Worksheet"
"Operational HPF-Worksheet"
"Criterion Worksheet"

The weights that you will give to nodes below the SPI level do not play any role at this stage, but they will be used in Chapter 7 for the Evaluation Tree.

Use the example in Figure H5-2 while you are completing the following instructions. On the copy of the Selection Tree that you prepared at the end of Chapter 4, do the following:

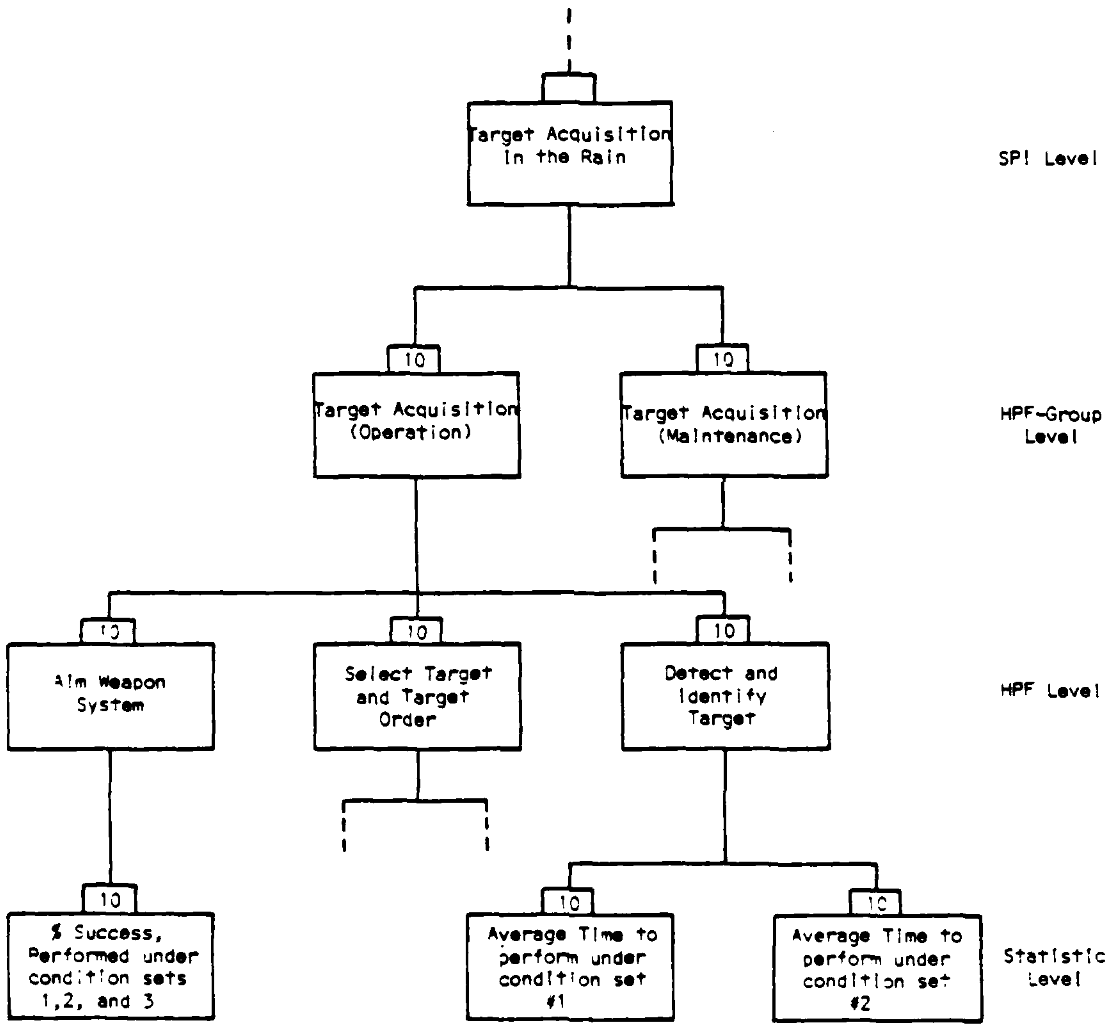
(1) Decompose each SPI on your Selection Tree to one or more nodes according to the number of HPF-Groups you have for this SPI. Record the HPF-Groups name in the appropriate space.

Usually each SPI will have 2 HPF-Groups associated with it: an Operational HPF-Group and a Maintenance HPF-Group.

(2) Decompose each HPF-Group you recorded in (1) to its selected HPF's.

The required information is recorded on the "Maintenance Worksheets" for Maintenance HPF-Groups. The selected HPF's for each Operational HPF-Group are recorded in the lefthand column of the "Final Condition Set Worksheets." These Worksheets are divided according to their SPI's and System Functions. For each HPF-Group, first draw a number of nodes corresponding to the number of HPF's in this HPF-Group. Record in each node the appropriate name of the HPF.

Figure H5-2
EXAMPLE SELECTION TREE



(3) Decompose each HPF you recorded in (2) to its statistic(s) taking into consideration the various condition sets under which the specific HPF is going to be tested.

The required information at this stage is recorded on the "Criterion Worksheet(s)." For each HPF, the number of statistics to be calculated depends on the different condition sets under which the HPF is to be tested and on the decision as to whether to calculate separate statistics for each condition set or one statistic across condition sets. As an aid in making this determination, the number of statistics for each HPF correspond to the number of different statistic criteria. For each HPF, first draw the number of nodes corresponding to the number of statistics. Record in each node the statistic type (percentage or average) and the condition set number(s) under which each statistic will be calculated. If a statistic is to be calculated across more then one condition set, its node should include the numbers of all the sets. If a statistic is to be calculated for one condition set, there should be only one number in the node.

(4) Attach weights to all nodes created in this chapter.

Following the SPI level, all nodes in the same level of the Tree should have the same weight. It is not important which weight you pick. For convenience of calculation we suggest the number 10. These numbers will be used in the Evaluation Tree discussed in Chapter 7.

A Technical Suggestion for Drawing the Selection Tree:

If the Selection Tree becomes so large it is not possible to contain it all, even on a reasonably large sheet of paper, divide the Tree in two. The first levels of the Tree which includes: System, System Functions, and SPI's would be drawn on <u>one</u> piece of paper. HPF-Groups and the levels underneath them would be drawn on separate sheets of papers for each SPI. Thus, each SPI and the nodes which branch from it would be described on a spearate piece of paper. Following this method, the number of pieces of paper needed to contain one Selection Tree would be equal to the number of SPI's + 1.

NOTE: Since you have a group of knowlegeable experts who developed the "Criterion Worksheets" it is recommended that you now perform the procedures in Chapter 7 section III which develops a Value Function for each statistic. The procedure in these sections should be performed by experts with a knowledge of the HRTES performance and statistic criteria.

6. HUMAN PERFORMANCE MEASUREMENT

CONTENTS:

ACTIONS:

(1) Determining the number of trials for each HPF in the field test.

(2) Planning and preparing for the collection of performance data.

(3) Preparing for collection of diagnostic data during the OT.

(4) Computing the statistics from field test data.

(5) Computing the confidence limits of the statistics.

PRODUCTS:

(1) Performance Data Collection Worksheets.

(2) Beginning and End Point Worksheets.

(3) Planning Data Collection Worksheet.

(4) OT Training Data Collection Worksheets.

(5) HPF Difficulty Worksheets for Observers and Participants and Performance Difficulty Questionnaires for Observers and Participants.

(6) Critical Incident Reports.

(7) Opinion Summary Data Worksheets.

(8) Statistic Worksheets.

USED FOR: (1) Working papers for the Outline Test Plan.

(2) Working papers for the Test Design Plan.

(3) Working papers for the Detailed Test Plan.

(4) Worksheets for direct inclusion in the Detailed Test Plan.

(5) Data for the Test Report.

(6) Data for the Independent Evaluation Report.

I. Determining the Number of Trials.

In planning an Operational Test, a stage is reached when a decision must be made about the number of trials to be performed for each of the HPF's. If the number of trials is too large, resources are being wasted; if it is too small, the reliability of the test results is significantly reduced. In this context, reliability is the extent to which test results are representative of those which would be obtained from the real population of users when the system is in the field.

In essence, there are two related decisions that must be made with regard to the field test. The first is the number of trials each performance unit will perform in the test. The second decision is the number of performance units which will take part in the test.

A performance unit consists of the individual(s) needed for a single system to perform the HPF. For example, the performance unit for firing a rifle consists of one person, whereas the performance unit for detecting targets in a medium tank may consist of two people (the commander and the gunner).

It is suggested that you try to get statistical assistance from experts regarding the number of trials per performance unit and the number of units needed to perform each HPF under each condition set. If statistical guidance is not available, use the following general guidelines to determine the number of trials per performance unit, and the number of performance units needed.

To calculate the number of trials and/or the number of performance units for the Operational Test, one has to determine two parameters:

(1) The maximum error permitted.

(2) The level of confidence that is acceptable.

Chances are that no sample which is taken will be an exact representation of the real population which will use the system in question. For this reason, the statistics which are taken on the sample population will probably provide somewhat different results than you would obtain by using the entire population. The difference between the results from the sample population and those from the entire population is the <u>error</u>. In general, as the size of your sample increases, and therefore becomes more like the real population, the error will decrease; however, your testing expenses will increase. For this reason, you have to decide on the maximum size of this error; that is, the difference between sample and entire population results which is acceptable. Suppose you have decided that you can permit a maximum error of five percent between sample population testing results and real population results; you then have to decide how much confidence you must have that your results will not exceed this error. This is the <u>level of confidence</u> you require. It is expressed as a percentage. For example, you may decide that you must be <u>90%</u> confident that your results will <u>not</u> exceed a five percent error. The higher the level of confidence which you find minimally acceptable, while holding the error constant, the larger the sample population will have to be. Consequently, you will need a larger number of trials per performance unit and a larger number of performance units in the Operational Test.

The following procedure applies to each statistic. Therefore, if a statistic is calculated for each condition set of an HPF separately, the number of trials and performance units applies to each condition set separately. If, on the other hand, you aggregated your statistic across condition sets, then the number of trials and performance units applies to all the condition sets <u>together</u>. In this case, it will be necessary to divide the number of trails equally between the condition sets which were aggregated.

(1) Decide on the number of trials per performance unit for each HPF under each condition set.

It has been determined, based on experimental statistical research, that at least <u>five trials for each performance unit</u> should be used in an Operational Test. This will supply the <u>minimum</u> amount of data necessary to meet the statistical assumptions underlying the determination of sample sizes. If less than five trials are used, the determination of sample size becomes more tenuous.

We will supply you with tables based on the assumption that you will use five trials per unit. However, we will also supply you with a formula to use if you decide on more or fewer than five trials per unit.

<u>Note</u>, it is not recommended that you use more than ten trials per unit.

(2) Decide on the number of performance units to be used to test each HPF under each condition set. This decision depends on whether the statistic used for this HPF is a percentage or an average. We will discuss each in turn.

<u>Percentage Statistics</u>.

(a) Decide which confidence level is acceptable. Conventionally, 95% is selected as the appropriate confidence level for rigorous experimentation, but in the case of Operational Testing, a lower confidence level may be acceptable. The lower the confidence level selected, the fewer performance units you will require.

Table H6-1 includes the 95%, 90%, 85%, and 80% confidence levels. No matter which confidence level you select, it must be the same for all statistics of all HPF's.

(b) Determine the maximum error, in terms of percentage points, that is acceptable. For example, if you accept an error of five percentage points, then, with your level of confidence, you could say that the real population value is within $\pm$ five percentage points of the result obtained during the Operational Test which used only a sample from that population. Of course, the larger the error you accept, the less meaningful your data will become. However, as the error permitted increases, the number of performance units required decreases.

(c) Table H6-1 indicates the number of performance units required, given various error and confidence levels. The table was constructed assuming five trials per performance unit, and the largest reasonable variance for the statistic. If fewer than five trials per unit are to be used, go to the next instruction to determine the number of units required.

Table H6-1
NUMBER OF PERFORMANCE UNITS REQUIRED

ERROR	CONFIDENCE LEVEL			
	95%	90%	85%	80%
5%	77	55	42	33
10%	20	14	11	8
15%	9	6	5	4
20%	5	4	3	2

(d) If five trials per performance unit are not used, the number of performance units required can be computed using the following formula.

Note: Do not use more than ten trials per unit.

$$N = \frac{k}{(\text{error}^2) \times n}$$

where:

N = Number of performance units required.

Error = Maximum error acceptable (in percentage points).

n = Number of trials per performance unit.

k = Is a constant which depends on the confidence limit, as follows:

CONFIDENCE LEVEL	k
95%	9604
90%	6806
85%	5184
80%	4096

Example: If there are ten trials per unit and you wish to be 90% confident that test results will be within $\pm$ 5 percentage points of the real population mean:

$$N = \frac{6806}{(5)^2 \times 10} = 27.2 \text{ therefore 28 performance units}$$

Average Statistics.

(e) Determine which confidence level is acceptable. See instruction (a) above for explanation.

(f) Determine the maximum error that is acceptable. For example, if you were measuring time to unload a truck, the maximum acceptable error might be 20 minutes. Be sure to express the error in the same units as the average (i.e., minutes, hours, etc.). With your level of confidence, you could then say that the real population average loading time would be within $\pm$ 20 minutes of the average obtained during the Operational Test.

(g) Estimate the maximum variance expected for a unit across its trials. A unit's performance will vary from trial to trial. The variance we are interested in is the variability of the most variable or erratic unit you foresee using in the Operational Test. Thus, variance can be estimated using one of the following three methods. These are listed in order of preference:

METHOD 1. Pilot Study.

Perform a pilot study with a representative unit and calculate the variance across trials using the following formula. A minimum of five trials should be used.

$$\text{Variance} = \frac{\Sigma X^2 - \frac{(\Sigma X)^2}{n}}{n-1}$$

where:

n = the number of trials performed by the unit.
X_i = the outcome of performance for trial i.

METHOD 2. Previous Data:

Estimate the variance of a unit across trials by examining data collected from a similar system, or an earlier OT of the same system.

METHOD 3. Expected Range:

Estimate the variance of a unit across trials by estimating the range of performance across trials and using the following formula. (Note: This formula assumes that performance of a unit is normally distributed across trials.)

$$\text{Variance} = \frac{(\text{Range})^2}{16}$$

where:

Range = highest expected value - lowest expected value

(h) Calculate the number of performance units required using the following formula:

$$N = \frac{kV}{n(\text{error})^2}$$

where:

N = Number of performance units required.

V = Maximum variance of a unit across trials.

error = Maximum acceptable error.

n = Number of trials per unit (minimum 5 recommended)

k = Constant whose value depends on confidence level selected:

CONFIDENCE LEVEL	k
95%	3.84
90%	2.72
85%	2.07
80%	1.64

Example: For time to load truck, using five trials per unit, and an estimated maximum variance of 100 minutes. If you wish to be 90% confident that the test mean will be within $\pm$ three minutes of the real population mean, then:

$$N = \frac{(2.72)\ (100)}{(5)\ (3^2)} = 6 \text{ performance units}$$

(3) Record the number of trials and number of units on the "Performance Data Collection Unit."

II. Planning Data Collection.

The planning of data collection is divided into the question of (a) what data are to be collected; (b) how are the data to be collected?

What Data to Collect?

All information required for this part is recorded on the "Criterion Worksheets" which were completed in Chapter 5. You will transfer the relevant information to the "Performance Data Collection Worksheets," in the form of raw data to be taken, and send them to the field test personnel for use in the OT.

(1) Copy a sufficient number of "Performance Data Collection Worksheets" (page W6-2) for one trial of each HPF.
(2) Fill in the top of each worksheet. (number of units)
(3) Fill in the relevant information from the "Criterion Worksheets." (See sample Worksheet, page H6-12.)

The information on the "Criterion Worksheets" was written in terms of errors. The information recorded in the "Data Collection Worksheet" should be specified in such a way as to call for the taking of raw data. These raw data may sometimes be in the form of errors, but frequently it will simply be the recorded performance. Later, this performance will be compared with test plan information to determine if an error has occurred. For example, if the "Criterion Worksheet" specifies that you want to know whether a target was not detected, the datum to be taken is the error. However, if "Criterion Worksheet" specifies that you want to know whether an error was made in identifying the model of the target, the datum to be taken is the actual model specified. This will prevent unnecessary loss of data. Later, field test personnel will compare the model designation with the actual model to determine if an error was made, and this information will be added in the 4th column of the worksheet. Additional checklists may be necessary to provide more space.

SAMPLE

HRTES PERFORMANCE DATA COLLECTION WORKSHEET

SYSTEM FUNCTION: DESTROY AIRCRAFT

SPI: ACQUIRE TARGET(S) IN FULL SUNLIGHT

HPF: DETECT AND IDENTIFY TARGETS

☒ OPER CONDITION SET(S): 3

☐ MAINT

ONE PERFORMANCE UNIT: ONE PARTICIPANT

NUMBER OF TRIALS PER PERFORMANCE UNIT RECOMMENDED: 5

NUMBER OF PERFORMANCE UNITS RECOMMENDED: 20 TRIAL # 1

PERFORMANCE UNIT DESIGNATION: ALPHA

DATA TO BE TAKEN*	OUTCOME	ACCIDENT OR NEAR ACCIDENT	SPECIFY IF UNSUCCESSFUL AND EXPLAIN REASON
SPECIFY IF TARGET SELECTED WHEN NO TARGET WAS PRESENTED			
— OR —			
SPECIFY IF TARGETS NOT DETECTED WHEN TARGET WAS PRESENTED			
— OR —			
SPECIFY WHAT WAS IDENTIFIED, FRIENDLY OR ENEMY			
— AND —			
SPECIFY MODEL OF TARGET CALLED OUT IN IDENTIFICATION			
— OR —			
SPECIFY TIME FROM TARGET PRESENTATION TO IDENTIFICATION			

*If time is to be measured, specify time units

SYSTEM MERCURY AIR DEFENSE WEAPON SYSTEM TEST OT-II DATE 25 MAR 81 PAGE ____

NAME ________________________ TELEPHONE ____________

(3) Make one copy of the filled-in Worksheet from (2) above for each trial to be run.

(4) Copy the relevant "Criterion Worksheets" completed previously in Chapter 5 and attach them to the appropriate "Performance Data Collection Worksheet."

(5) Hold the completed Worksheets. They will be attached to the "Beginning and End Points for Time Measures Worksheets."

<u>How are the Data to be Collected?</u>

To plan data collection one has to decide whether required data will be taken by observers, instrumentation, or a combination of the two. HRTES includes "Planning Data Collection Worksheet" (pages W6-4 - W6-11). You may use it as a guide or actually copy it and answer its questions. It is designed to aid you in considering how data are to be taken and what is required to prepare for data collection. The information developed using this worksheet should be included in the "Detailed Test Plan."

In addition, if a time measure is involved, it is necessary to analyze each HPF chosen for evaluation to identify the <u>beginning</u> and <u>end</u> points of performance. The following instructions will aid you in this identification:

(1) Copy a sufficient number of "Beginning and End Point Worksheets." Worksheets," (page W6-3).

(2) Determine the unit of measure (i.e., fraction of a second, seconds, minutes, hours, etc.). The shorter the expected duration of performance, the smaller your unit of time is likely to be. The following are generally applicable:

(a) Performance which consists entirely of reaction time may usually be adequately measured in units of tenths of a second.

(b) For HPF's which have expected performance time in hours, the units of measure is minutes; for those with expected performance time in minutes, the unit of measure is seconds, etc.

(3) Identify the cue which initiates performance.

(a) For the first HPF of a Group, this cue is normally the introduction of a new stimulus such as: visual target, auditory signal, command, an internal mental process or decision.

(b) For later HPF's in a Group, the cue is often the completion of the previous HPF. It should be noted that in some cases several HPF's may be performed at the same time.

(c) It is possible that the initiating cue will not be directly detected by observers or instrumentation. In this case, you will have to specify an artificially-introduced cue which informs observers or instrumentation that performance has begun. The decision to introduce such signals should be made in advance, and indicated in the test plans.

(4) Identify the end cue, which informs OT personnel or instrumentation that performance has been completed, or that it has ceased in the absence of completion:

(a) The end cue is usually identifiable if there is resultant output. When this is not the case, system operators/ maintainers may have to signal their completion. The decision to introduce such signals should be made in advance and indicated in test plans.

(b) In the case of cessation of performance in the absence of completion, some clear signal must be given by the performer, or careful observation must take place. In the confusion which often accompanies this situation, the end point may not be clear. For this reason, subjects and observers must be trained, in advance, to signal the point at which the attempt to perform has ceased without completion.

(5) Record the beginning and end points of performance for HPF's on the "Beginning and End Points Worksheet," page W6-3. (See sample "Beginning and End Point Worksheet," page H6-16).

(6) Attach the completed worksheets to the appropriate "Performance Data Collection Worksheets" to be sent to field test personnel.

III. Collecting Diagnosis Data During the OT.

Certain data have to be collected to diagnose the probable cause(s) of sub-criterion HPF performance. A reasonable amount of these data can be collected after the conclusion of the field test. Some of it must be collected both before the start of the field test and while the field test is being run. There are two types of data to be collected: training data from the OT trainers, and opinion data from the OT participants and observers. Both types of data should be taken as soon as possible following the action to be measured. HRTES provides you with two types of Worksheets for taking these data.

SAMPLE

HRTES

BEGINNING AND END POINT WORKSHEET

SYSTEM FUNCTION: DESTROY AIRCRAFT

SPI: ACQUIRE TARGET(S) IN FULL SUNLIGHT

HPF-GROUP: TARGET ACQUISITION

☒ OPER
☐ MAINT

CONDITION SET: # 1

HPF's	BEGINNING POINT	END POINT	UNIT
DETECT AND IDENTIFY TARGET (S)	INTRODUCTION OF FIRST TARGET	IDENTIFICATION SIGNAL	1/10 SEC.
SELECT AND ORDER TARGETS	IDENTIFICATION OF ALL TARGETS	ORDERING COMPLETION SIGNAL	1/10 SEC.
ORIENT WEAPON SYSTEM	START OF ORIENTING	COMPLETION OF ORIENTING	SECS.
DETERMINE RANGE OF TARGET	DETECTION SIGNAL	RANGE CALL-OUT	1/10 SEC.
SHIFT TO SECOND TARGET	FIRING AT FIRST TARGET	COMPLETION OF SHIFT	SECS.

SYSTEM MERCURY AIR DEFENSE WEAPON SYSTEM TEST OT-II DATE 25 MAR 81 PAGE ____

NAME ____________________ TELEPHONE ____________________

Training Data.

(1) Make one copy of the "Guidelines for the OT Training Data Collection Worksheet" (page W6-12) for each OT trainer.

(2) Make copies of the "OT Training Data Collection Worksheet" (page W6-13) for each OT trainer.

(3) Make a list of all HPF's to be trained and copy it for each OT trainer.

(4) Combine the Worksheets, Guidelines and HPF Lists into sets for each OT trainer.

(5) Submit these sets to the OT trainers prior to OT training. They are to be completed and returned to the OT field test managers immediately following training.

Opinion Data.

(1) Make one copy of "Guidelines for Collecting Data for Diagnosis During OT" (page W6-14) for each observer.

(2) Make at least one copy of the "HPF Difficulty Worksheet for OT Observers" and the "HPF Difficulty Worksheet for OT Participants" (pages W6-16 and W6-17) for each group of HPF's to be performed together. Usually these will be the HPF's in a given HPF-Group.

(3) Fill-in the required information at the top of the Worksheet, and the names of the HPF's to be performed together.

(4) If there will be more than one observer or participant for a given HPF, make the appropriate number of copies of the Worksheets you filled in.

(5) Make sufficient copies of the "Performance Difficulty Questionnaire for Participants" and the "Performance Difficulty Questionnaire for Observers." These Questionnaires start on pages W6-18 and W6-27.

(6) Make sufficient copies of the "Critical Incident Report" (page W6-38) for each OT observer and participant.

(7) Submit the resulting packages of data collection instruments to the appropriate field test managerial personnel with instructions that they are to be filled in, as required, following the completion of the testing of each group of HPF's. These packages will consist of:

HPF Difficulty Worksheet for OT Observers
HPF Difficulty Worksheet for OT Participants
Guidelines for Collecting Data for Diagnosis During OT
Performance Difficulty Questionnaire for Participants
Performance Difficulty Questionnaire for Observers
Critical Incident Report

(8) After you have received the <u>completed</u> data collection instruments from the field test personnel, copy the "Opinion Summary Data Worksheet" (page W6-39). Make enough copies so that you have one for each HPF scored 50 or above.

(9) Write in the name of each HPF tested, its SPI, System Function, and condition set on a separate "Opinion Summary Data Worksheet." (See sample Worksheet, page H6-20.)

(10) Compute the overall mean of each opinion scale, for observers and participants who rated the HPF 50 or above in difficulty.

(11) Record the computed means, for a given HPF, on the appropriate "Opinion Summary Data Worksheet."

(12) All "HPF Selection Worksheets" should be retained, as their "Comments" section may prove useful later in the diagnostic phase of the OT.

IV. Processing Performance Data from the Field.

After the field test is completed, you will be provided with the data collected. These data have been recorded on the "Performance Data Collection Worksheets." Each such Worksheet will provide you with raw data for each trial of each HPF and will indicate if the outcomes were a failure or a success. You will be asked to process these data on the "Statistic Worksheets." Each of these Worksheets will contain data on all trials for all performance units of each HPF. For each HPF, you will then calculate the appropriate statistics and the confidence limits.

(1) Make at least one copy of the "Statistic Worksheet" (page W6-40) for each HPF tested.

(2) Fill in the required information at the top of each Worksheet. This information can be copied from the appropriate "Criterion Worksheets." (See sample Worksheets, pages H6-20 and H6-21 .)

(3) Each "Performance Data Collection Worksheet" contains the data required to fill in one cell of the Trials X Units Matrix of the "Statistic Worksheet" for the given HPF.

SAMPLE

HRTES STATISTIC WORKSHEET

SYSTEM FUNCTION: DESTROY AIRCRAFT

SPI: ACQUIRE TARGET(S) IN FULL SUN LIGHT

HPF: DETECT AND IDENTIFY TARGETS

CONDITION SET(S): 1 AND 3

☒ OPER
☐ MAINT

TYPE OF STATISTIC: PERCENTAGE

UNITS

TRIALS	1	2	3	4	5	6	7	8	9	10	
1	⊖	⊖	/	/	⊖	⊖					
2	/	⊖	⊖	⊖	/	⊖					
3	/	/	/	/	/	/					
4	/	/	/	/	⊖	/					
5	/	/	/	/	/	⊖					
6											
7											
8											
9											
10											
SUMS	4	3	4	4	3	2					GRAND SUM
SUMS/TRIALS	.80	.60	.80	.80	.60	.40					.4

STATISTIC OUTCOME: 67%

LEVEL OF CONFIDENCE: 95%

UPPER CONFIDENCE LIMIT: 84%

LOWER CONFIDENCE LIMIT: 50%

SYSTEM MERCURY AIR DEFENSE WEAPON SYSTEM TEST OT-II DATE 25 MAR 81 PAGE ____

NAME ____ TELEPHONE ____

SAMPLE

HRTES

STATISTIC WORKSHEET

SYSTEM FUNCTION: DESTROY AIRCRAFT

SPI: ACQUIRE TARGET(S) IN FULL SUNLIGHT

HPF: AIM SYSTEM

CONDITION SET(S): 3

☒ OPER
☐ MAINT

TYPE OF STATISTIC: AVERAGE (DEVIATIONS IN METERS)

UNITS

TRIALS	1	2	3	4	5	6	7	8	9	10	
1	20	50	10	25	40	20					
2	25	50	10	30	20	20					
3	15	40	5	20	10	20					
4	15	50	~~0~~	20	5	15					
5	10	30	5	10	10	15					
6											
7											
8											
9											
10											
SUMS	85	220	30	105	85	90					GRAND SUM
SUMS/TRIALS	17	44	6	21	17	18					123

STATISTIC OUTCOME: 20.5 METERS

LEVEL OF CONFIDENCE: 95 %

UPPER CONFIDENCE LIMIT: 23.38 METERS

LOWER CONFIDENCE LIMIT: 17.62 METERS

SYSTEM MERCURY AIR DEFENSE WEAPON SYSTEM TEST OT-II DATE 25 MAR 81 PAGE ____
NAME ______________________ TELEPHONE ______________

If the statistic is a percentage, write 1 (Success) or 0 (Failure) in the appropriate cells of the matrix. This information will be found in the righthand column of the "Performance Data Collection Worksheet."

If the statistic is an average, there will only be one outcome listed. Copy this outcome in the appropriate cell of the Worksheet. All trials of all performance units of this HPF, under a given condition set, will be recorded on this Worksheet.

(4) Calculate the statistic for each HPF, on each "Statistic Worksheet." Whether your statistic is a percentage or an average do the following:

(a) Sum the numbers in each column (units), and record each sum in the "Sums" row.

(b) Divide each sum by the number of trials it included, and record the resulting quotients in the "Sums/Trials" row.

(c) Sum all the quotients in the "Sums/Trials" row and record it in the "Grand Sum" cell.

(d) Divide the Grand Sum by the number of units. If your statistic is an average, record this number in the "Statistic Outcome" space. If your statistic is a percentage, multiply the result of the division above by 100, and record it in the "Statistic Outcome" space.

The formulas for the calculation of the statistics you have computed are:

$$\text{Average} = \frac{\sum_{i=1}^{N} \left(\frac{\sum_{j=1}^{n} x_{ij}}{n} \right)}{N}$$

$$\text{Percentage}=100 \left(\frac{\sum_{i=1}^{N} \left(\frac{\sum_{j=1}^{n} x_{ij}}{n} \right)}{N} \right)$$

where:

x_{ij} = Cell for unit i, trial j.

n = Number of trials per units.

N = Number of performance units.

(5) Calculate the upper and lower confidence limits of each statistic and record them in the appropriate spaces on the Worksheet. These calculations may be done by statistician, otherwise perform following procedure.

(a) If the statistic is a <u>percentage</u>, the formulas for upper and lower confidence limits are as follows: (For the "upper" use the "+" sign in the formula; for the "lower" use the "-" sign in the formula.)

$$\text{Percentage outcome} \pm 100k \sqrt{\frac{\sum_{i=1}^{N} P_i(1-P_i)}{N^2 n}}$$

where:

N = Number of units.

n = Number of trials per unit.

k depends upon the level of confidence required:

CONFIDENCE LEVEL	k
95%	1.96
90%	1.64
85%	1.44
80%	1.28

P_i = the proportion of success for unit i. These numbers were already computed and recorded on the sums/trials row of the Worksheet. (See the following example.)

Example: Based on the data in the "Sample Statistic Worksheet" (page H6-21), the confidence limits for 95% level of confidence are:

$$67 \pm 1.96 \times 100 \times \sqrt{\frac{0.8(1-.08)+0.6(1-0.6)+0.8(1-0.8)+0.8(1-0.8)+0.6(1-0.6)+0.4(1-0.4)}{36\times 5}}$$

$$= 67 \pm 1.96 \times \sqrt{\frac{1.36}{36\times 5}} = 67\% \pm 17\%$$

thus the upper limit is 84%

the lower limit is 50%

(b) If your statistic is an <u>average</u> the formulas for upper and lower confidence limits are as follows:

$$\text{Average outcome} \pm k\sqrt{\frac{\sum_{i=1}^{N} V_i}{N^2 n}}$$

where:

N, n, k, are as in the case of percentages. V_i is the variance of performance in unit i. It can be computed according to the following formula:

$$V_i = \frac{\sum_{j=1}^{n} X_{ij}^{2} - \frac{(\sum_{j=1}^{n} X_{ij})^2}{n}}{n-1}$$

where:

X_{ij} = the content of the cell for unit i, trial j.

Example: Based on the data in the "Sample Statistic Worksheet" (page H6-22) the confidence limits for 95% level of confidence are:

$$20.5 \pm 1.96\sqrt{\frac{\sum_{i=1}^{6} V_i}{36x5}}$$

where:

$$V_1 = \frac{20^2+25^2+15^2+15^2+10^2 - \frac{85^2}{5}}{4} = 32.5$$

$$V_2 = \frac{50^2+50^2+40^2+50^2+30^2 - \frac{220^2}{5}}{4} = 80$$

$$V_3 = \frac{10^2+10^2+5^2+0^2+5^2 - \frac{30^2}{5}}{4} = 17.5$$

$$V_4 = \frac{25^2+30^2+20^2+20^2+10^2 - \frac{105^2}{5}}{4} = 55$$

$$V_5 = \frac{40^2+20^2+10^2+5^2+10^2 - \frac{85^2}{5}}{4} = 195$$

$$V_6 = \frac{20^2+20^2+20^2+15^2+15^2 - \frac{90^2}{5}}{4} = 7.5$$

Substituting the outcomes for the various variances we get:

$$20.5 \pm 1.96 \sqrt{\frac{32.5+80+17.5+55+195+7.5}{36 \times 5}}$$

$= 20.5$ meters $\pm$ 2.88 meters

Thus the upper limit is 23.38 meters and the lower limit is 17.62 meters.

7. EVALUATION

CONTENTS:

I. Introduction.

II. Building the Evaluation Tree.

III. Creating Value Functions.

IV. Evaluating Results after OT Data are Obtained.

V. Interpreting the Evaluation Tree.

VI. Answering Issues that Do Not Directly Question System Performance.

ACTIONS:

(1) Converting the basic structure of the Selection Tree to the structure of the Evaluation Tree.

(2) Deriving evaluation weights for each node of the Evaluation Tree.

(3) Creating a Value Function for each statistic of each HPF.

(4) Following the field test, determining the value of each statistic's outcome and confidence limits.

(5) Interpreting the meaning of the completed Evaluation Tree.

PRODUCTS:

(1) A Value Function for each statistic of each HPF.

(2) A completed Evaluation Tree.

(3) An interpretation of the completed Evaluation Tree.

USED FOR:

(1) Working papers for Detailed Test Plan (if interactive Plan is written).

(2) Working papers for Independent Evaluation Plan.

(3) Evaluation Tree for direct inclusion in Independent Evaluation Plan.

I. Introduction.

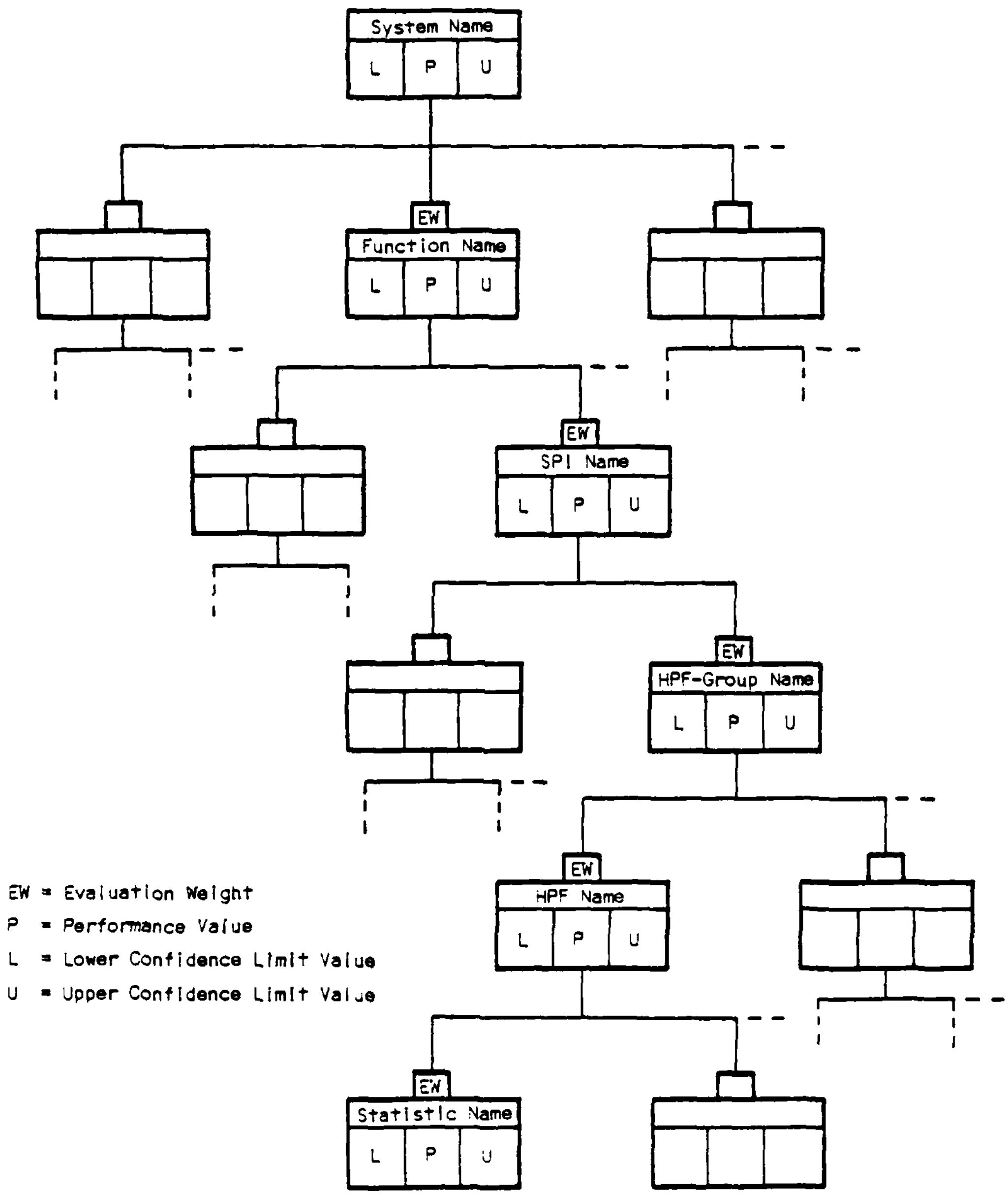
The basic method for evaluating performance in the OT involves creating a logical hierarchical structure. This structure is called an Evaluation Tree. The Evaluation Tree for HRTES consists of the following hierarchy, as viewed from the top down:

(1) The System
(2) System Functions
(3) System Performance Issues (SPI's)
(4) Human Performance Function Groups (HPF-Groups)
(5) Human Performance Functions (HPF's)
(6) Statistics

Each hierarchical level of the Evaluation Tree contains one or more elements called <u>nodes</u>. That is, the System node branches into the nodes for the System Functions. The nodes for the System Function branch into those for their relevant SPI's. The nodes for the SPI's branch into those for their relevant HPF-Groups, etc.

Each node is given a weight which indicates its criticality, relative to all other nodes which branch from a common node in the preceeding level. That is, all SPI's which branch from a single System Function are weighted to indicate their relative criticality for that System Function, etc. Human performance statistics occupy the bottom level of the Evaluation Tree. Figure H7-1 illustrates the general structure of the Evaluation Tree. The performance value (P), lower confidence limit value (L), and upper confidence limit value (U) shown in Figure H7-1 are described in section III of this chapter.

Figure H7-1
EVALUATION TREE STRUCTURE



To evaluate the outcomes of the human performance statistics, approximated Value Functions are developed and the actual outcome of each statistic is then assigned a value. Assigning outcome values allows you to compare heterogeneous measures of human performance. Once you have developed approximated Value Functions with the aid of experts, and after all the data are collected, it is possible to derive values of performances for each node in the Tree through the process of folding back the Evaluation Tree. By this method, it is possible to evaluate each HPF, HPF-Group, SPI, and System Function, as well as the whole System.

II. Building the Evaluation Tree.

The major work of building the Evaluation Tree has already been done in the process of building the Selection Tree. The Evaluation Tree has basically the structure of the Selection Tree with the omission of those nodes that were not actually tested in the OT. In addition, each node in the Evaluation Tree *will* contain some pieces of information which are needed for evaluation. You will be asked to perform some alterations in the Selection Tree. When these alteration are complete, the Selection Tree will have become the Evaluation Tree.

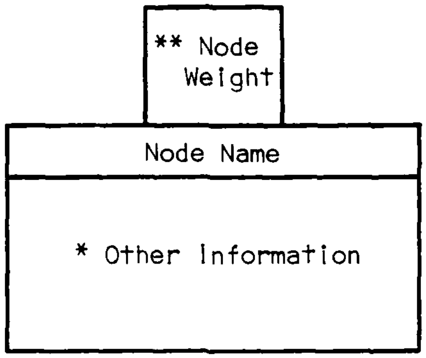
The following instructions consist of two phases: Creating the Tree structure, and Deriving Evaluation Weights.

Creating the Tree Structure:

After the field test has been performed:

(1) Indicate on your copy of the "Selection Tree," whether or not each of the nodes was actually tested. If you are building the Evaulation Tree before the field test, ignore this instruction.

(2) Copy the structure of the Selection Tree on one large sheet of paper, making sure to eliminate all those nodes which were not tested in the OT, if this is applicable. For the Evaluation Tree we suggest the following format for a node:



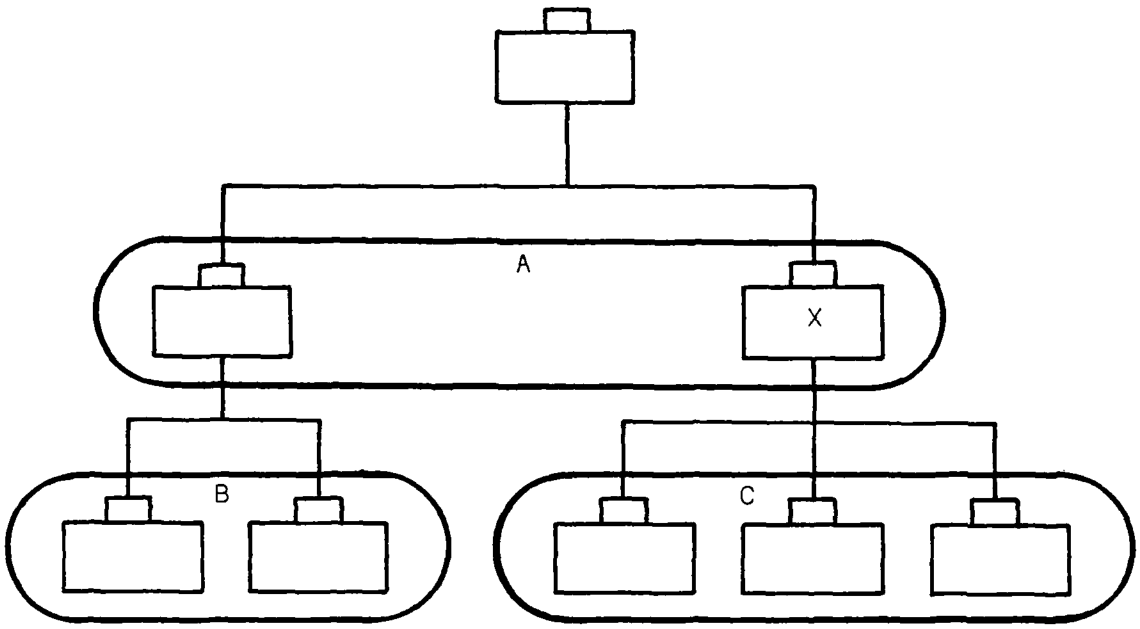
*other information - will be specified later
**for the System Node you will omit node weight

(3) Copy the name of the node from the Selection Tree, and record it in the appropriate place in the Evaluation Tree.

To save space, it is recommended that you not write the node's type (i.e., SPI, HPF, etc.). The type of node will be clear from its level in the Tree.

Deriving Evaluation Weight

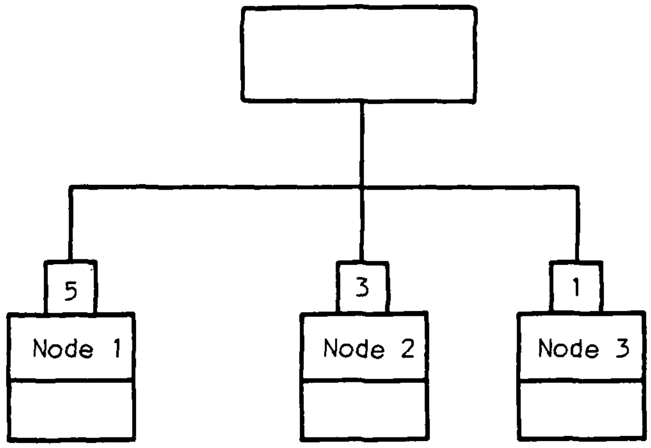
Examine the Evaluation Tree you have created so far. You will notice that in each level the nodes can be divided into groups which will be called families. Each family is a set of nodes which branch from the same node in the above level. This node is called the "parent" of the family. In the figure on page H7-6 , families are marked by circles. Node x is the "parent" of the family C and belongs to family A.



The evaluation weight of a node is based on its selection weight and the selection weights of all nodes in its family. The evaluation weight of a node is calculated by dividing its selection weight by the sum of selection weights of all nodes in its family, as follows:

$$\text{Evaluation Weight of a Node} = \frac{\text{Selection Weight of the Node}}{\text{Sum of Selection Weights of all Nodes in it's Family.}}$$

For example, given the following family of nodes from the Selection Tree, with the Selection Weights 5, 3, and 1 as shown:



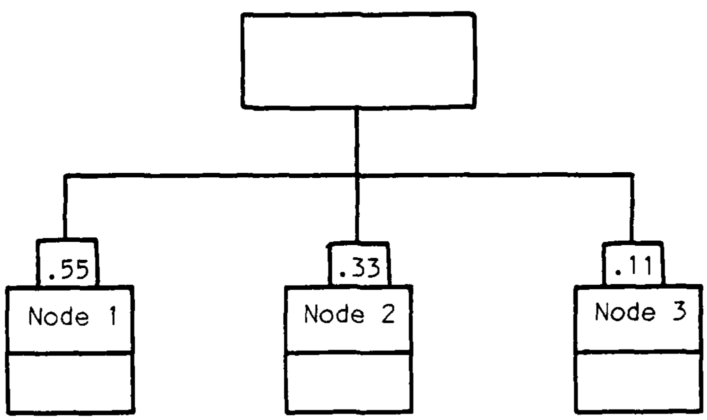
The Evaluation Weights are calculated as follows (where E.W. = Evaluation Weight):

$$\text{E. W. of node 1} = \frac{5}{5+3+1} = \frac{5}{9} = .55$$

$$\text{E. W. of node 2} = \frac{3}{5+3+1} = \frac{3}{9} = .33$$

$$\text{E.W. of node 3} = \frac{1}{5+3+1} = \frac{1}{9} = .11$$

Thus, in the Evaluation Tree you will have the following weights:



This alteration of weights is done to make the weights of a family of nodes sum to 1. Instructions for deriving evaluation weights follow:

For each family of nodes do the following:

(1) Sum the selection weights of nodes in the family. Remember: Do not include weights of nodes which were not tested and thus not included in the Evaluation Tree.

(2) Derive the evaluation weights for each node in the family by dividing its selection weight by the sum of its family from (1).

(3) Record the evaluation weights on the available spaces in the Evaluation Tree.

III. Creating the Value Functions.

For each statistic, whether it is an average type or percentage type, a function must be found that will yield a value for any outcome of that statistic. By means of this Value Function it will be possible to estimate the value of any outcome of a statistic. The Value Function will convert the outcome of differing statistics to a common scale. All Value Functions will range between zero and one hundred. The Value Function Scale is as follows:

Very Bad	Bad	Borderline	Good	Very Good
0	25	50	75	100

This procedure should be performed by the same experts who developed the "Criterion Worksheets" in Chapter 5 immediately following development of the Criterion Worksheets.

(1) For each expert:

(a) Copy the "Guidelines for Defining Value Functions" (page W7-2 through W7-7).

(b) Copy one "Value Function Worksheet" (page W7-8) for each statistic which is to be taken during the OT.

(c) Fill in the required information at the top of each "Value Function Worksheet," then attach the appropriate "Criterion Worksheet" and "Final Conditions Set Worksheet."

(d) Copy the "Sample Value Function Worksheet" (page W7-9).

(2) Submit the above items to each expert.

(3) After you have received the completed "Value Function Worksheets" for each statistic from the experts, compute the mean of the "very good" outcomes and the mean of the "very bad" outcomes for each statistic.

(4) On a fresh Value Function Worksheet enter the mean "very good" and mean "very bad" outcomes in the appropriate boxes on the graph. (See "Sample Value Function Worksheet," page H7-10).

(5) Place the statistic criterion of this statistic in the appropriate position between the "very good" outcome and the "very bad" outcome in the bottom scale of the graph. If you have no criterion for this statistic and are unable to obtain one, then no statistic criterion need be entered on the graph.

SAMPLE

HRTES VALUE FUNCTION WORKSHEET

SYSTEM FUNCTION: DESTROY AIRCRAFT

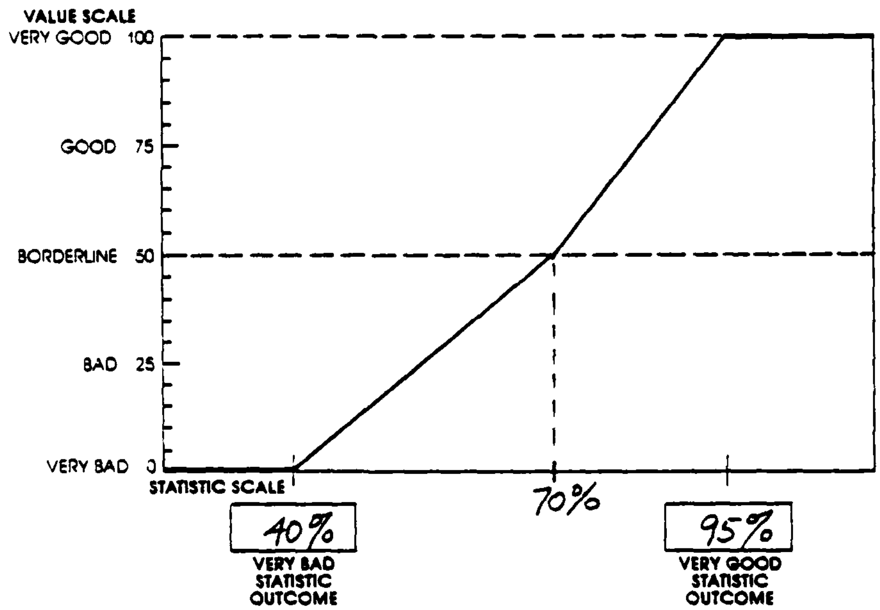
SPI: ACQUIRE TARGET(S) IN FULL SUNLIGHT

HPF: DETECT AND IDENTIFY TARGETS

CONDITION SET: 3

☒ OPER.
☐ MAINT.

STATISTIC: PERCENTAGE



STATISTIC CRITERION: 70 %

SYSTEM MERCURY AIR DEFENSE WEAPON SY. TEST OT-II DATE 25 MAR 81 PAGE ____

NAME ____________________ TELEPHONE ____________________

(6) Plot the Value Function for the statistic by connecting the three points: "very bad" (value - 0), criterion (value = 50), and "very good" (value = 100) by straight lines. If no criterion is available you will connect the "very good" and "very bad" points to generate the Value Function. NOTE: If you did not have a criterion, the default statistic criterion will be generated by your Value Function for the statistic in question. It will be halfway between the "very good" and "very bad" points.

IV. Evaluating Results after OT Data are Obtained.

When the field test is completed, you will receive the data which were taken by the field test personnel. In Chapter 6, you processed the performance data. You calculated statistic outcomes, and confidence limits for each of the HPF's under various condition sets. The results were recorded on the "Statistic Worksheets."

Your role now is to use the Evaluation Tree, the Value Functions, and the actual outcomes from the field test to evaluate the human performance in the system.

Filling in the Statistic Level of the Evaluation Tree

For each statistic and its appropriate Value Function:

(1) Find the value which corresponds to the statistic outcome (follow example in Figure H7-2, page H7-13).

(a) On the "Value Function Worksheets," place the actual statistic outcome (recorded on the "Statistic Worksheet") in the appropriate position between the "very good" outcome and the "very bad" outcome.

(b) Draw a vertical line from the position of the actual statistic outcome to the Value Function line.

(c) Draw a horizontal line from the resulting intertersection point on the Value Function line to the Value Scale. The intersection point on the Value Scale is the value of the statistic outcome.

(2) Find the values which correspond to the lower and upper confidence limits found on the "Statistic Worksheets."

This is done by performing steps a, b, c of (1) above for the confidence limits instead of the statistic outcome.

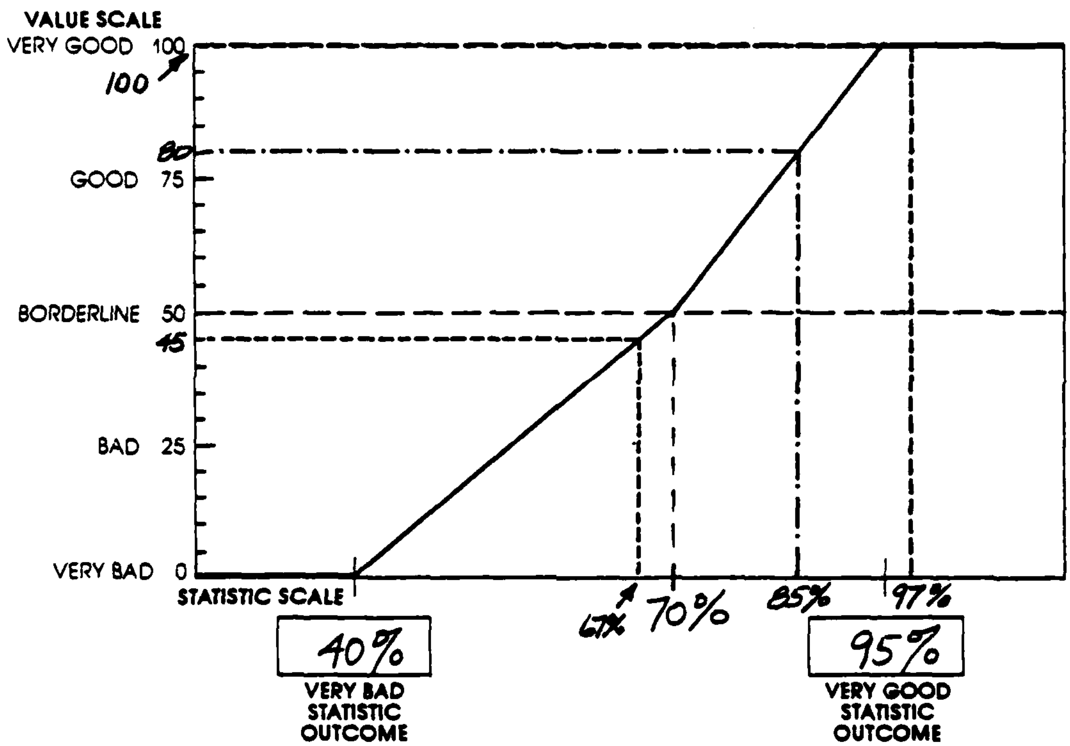
Example:

Assuming the statistic outcome for "Target Detection and Identification" was 85% and the lower and upper limits of this outcome were 67% and 97%, respectively, the values of this number derived from the Value Function graph as shown in Figure H7-2.

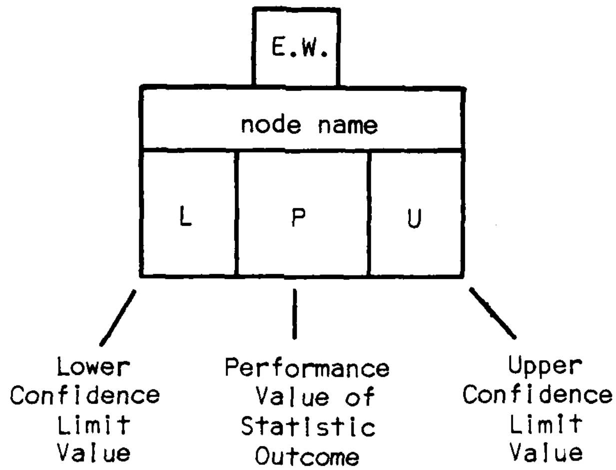
Thus, the values are:

	Value
Outcome = 85 yields	80
Upper Limit = 97 yields	100
Lower Limit = 67 yields	45

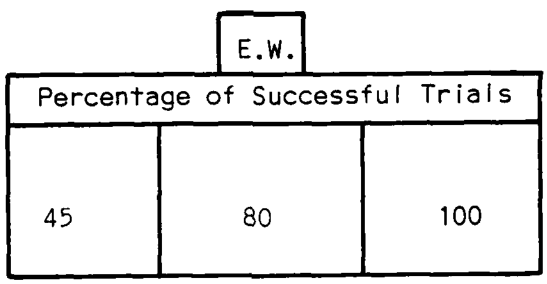
Figure H7-2
EXAMPLE UPPER AND LOWER
LIMITS GRAPHIC CALCULATIONS



(3) For each Statistic record the three values you derived in steps (1) and (2) on the appropriate Statistic Nodes of the Evaluation Tree as follows:



For the node corresponding to our example you would write:



At this point you have filled all the space for "other information" for the Statistic Nodes at the bottom of the Tree. In the following steps, you will be guided in filling in the "other information" space in all the nodes in the levels above the Statistic Level.

Folding Back the Tree

By folding back the Evaluation Tree all the "other information" space in the remaining nodes will be filled. In general, when you have the 3 values for each member in a family of nodes, you can calculate the 3 values of their "parent." Thus, the calculation proceeds from the lowest level to the level above it. Only when the calculation of one level of nodes is completed can you then continue with the calculations for nodes at the next higher level. The basic procedure which is performed recurrently is a weighted summation according to the formula:

$$P = \sum_{i=1}^{n} E.W._i \times P_i$$

Where:

P = Performance value of parent node

i = $i^{\underline{th}}$ node in the family

$E.W._i$ = Evaluation weight of node i in the family

P_i = Performance value of node i in the family

n = Number of nodes in the family

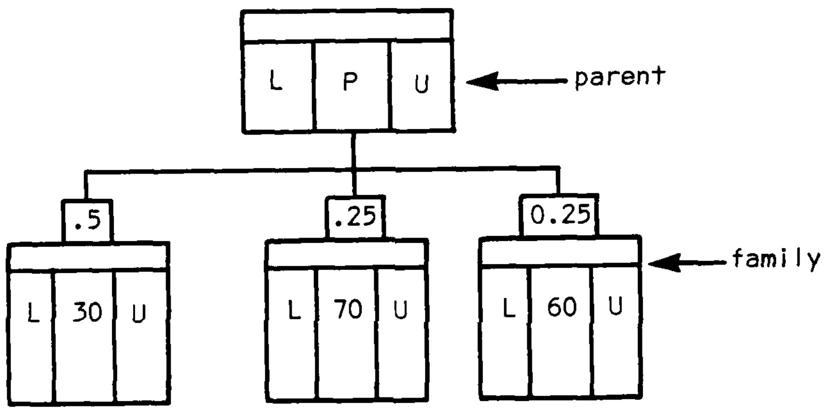
Example:

Given the performance values of a family of nodes, you can calculate the performance value of its parent as follows:

P = Performance value

L = Lower Confidence limit value

U = Upper Confidence limit value



The number which will be substituted for P in the parent node is:

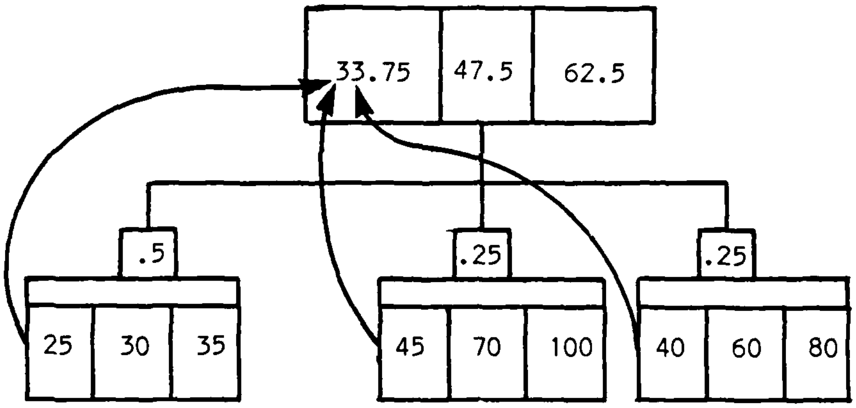
$$P = .5 \times 30 + .25 \times 70 + .25 \times 60 = 47.5$$

For deriving L or U in the parent node you will follow the same procedure. However, instead of using the P values (30, 70, 60) you will use the L (25, 45, 40) or U (35, 100, 80) values accordingly.

$$L = .50 \times 25 + .25 \times 45 + .25 \times 40 = 33.75$$

$$U = .50 \times 35 + .25 \times 100 + .25 \times 80 = 62.5$$

Thus, the performance values, and the upper and lower confidence limits for the parent node are as shown:



To fold back the tree for the performance values, perform the following steps. Refer to the example above for assistance.

(1) Multiply the P value of each Statistic node by its weight.

(2) For each family at the Statistic Level, sum the numbers you computed in step (1).

(3) Write the resulting sums in the "P space" of the appropriate parent HPF nodes. Note: If there is only one statistic for an HPF, the L, P, U values of the statistic can be copied in the L, P, U spaces of the parent HPF.

(4) Repeat these three steps for each successively higher level (HPF, HPF-Group, SPI, System Function, and finally System) in the Evaluation Tree.

(5) Repeat steps (1) through (4) for the upper (U) and lower (L) confidence limit values. Remember, for each node the following notation is the same:

P = Performance value
L = Lower confidence limit value
U = Upper confidence limit value

V. Interpreting the Evaluation Tree.

Interpreting the Evaluation Tree consist of analyzing the performance values of each of the nodes.

Each node contains three numbers: the value of lower confidence limit (L), the value of performance (P) and the value of upper confidence limit (U).

The higher these numbers are for a given node, the better is the performance at this node. It is possible to state that the "real" value of performance of a node is somewhere between the L number of this node and the U number of this node. By "real value" we mean the value which is determined based on performance of the whole population. The confidence in this statement is the level of confidence chosen.

For example: suppose you have chosen a confidence level of 90% for calculation and it turned out that, for a System Function, the values 70, 80, 90 were obtained for L, P, U accordingly. In this case, you can say with 90% confidence that the value of performance falls between 70 and 90.

The value numbers themselves can be interpreted by the Value Scale:

Very Bad	Bad	Borderline	Good	Very Good
0	25	50	75	100

Note: The smaller the interval between L and U the more meaningful are your results. For example, there is not much meaning to a statement such as "the value of performance is between "bad" and "good", i.e., between 25 and 75." There is a tradeoff between the level of confidence chosen and the confidence interval. If you accept a lower level of confidence, *the interval between L and U will be smaller.* Thus, you mav say that you are 90% confident that the value is between 70 and 90; whereas, you are 80% confident that the value is between 75 and 85, etc.

Through the process of developing Value Functions, the value of each criterion was always set at 50. It is possible now to use this criterion value to assess whether the performance for a node is acceptable. There are three possible situations:

(a) The L value of a node is greater than 50: In this case, you can state that the performance is acceptable at the confidence level selected.

(b) The U value of a node is less than 50: In this case, you can state that performance of the node is unacceptable at the confidence level selected.

(c) If 50 falls between L and U you will not be able to make any statement about the acceptability of the performance at the level of confidence chosen. However, you might be able to say something at a lower level of confidence if this is desirable. To do so, using a different level

of confidence, you will have to recompute the confidence limits, their values, and to fold back the new L and U numbers through the Tree.

VI. Answering Issues which Do Not Directly Question System Performance.
It is possible that issues from existing test planning documents, which do not directly question system performance, were included in the "Detailed Test Plan". Such non-performance issues question the effectiveness of: maintainability, safety, human factors engineering (HFE), training and user personnel, in the context of the system being tested.

Dealing with Effectiveness of Maintainability as an Issue

Each System Performance Issue (SPI) that was actually tested forms a node in the Evaluation Tree. Each SPI Node branches into parallel HPF-Group Nodes--for operational HPF's and maintenance HPF's. In this way, the value of each node, from the SPI Level of the Evaluation Tree and above, half derives from operational performance measurement and half derives from maintenance performance measurement.

If you wish to answer a specific issue of maintainability effectiveness, you simply fold back the Evaluation Tree using only the maintenance HPF-Groups, eliminating the operational HPF-Groups. This will give you the maintenance performance value for the system as a whole, for each System Function, for each maintenance HPF-Group, and for each maintenance HPF. After folding back the Evaluation Tree using only maintenance values you can apply *Section V. Interpreting the Evaluation Tree and thus answer any specific* maintenance effectiveness issue. In addition if you wish to address your evaluation to performance only (eliminating maintenance considerations), you can perform this same procedure using operational HPF-Group values and eliminating maintenance HPF-Group values.

Dealing with Effectiveness of Human Resource Areas as Issues - Introduction

The following is a description of techniques for answering human resource area issues for the system being tested. Normally such issues are concerned

with the effectiveness of: safety, human factors engineering (HFE), training, and user personnel. In HRTES safety is a component of HFE. However if required it can be broken out to answer a specific safety issue.

Before describing techniques for answering issues of overall effectiveness of human resource areas, a statement of the basic HRTES position on human resource areas is useful. Human resource area measurement and evaluation is significant for diagnosing the reasons for inadequate performance and/or for determining how to improve performance. Satisfactory human resource area design is important in so far as it leads to satisfactory performance over time. It does not have value if it is totally independent of performance. Performance, which combines system operation and system maintenance tasks required for the performance of all system missions, is a very general concept. It is so general as to be of questionable utility when used to define human resource design adequacy. To make human resource area measurements and evaluation as useful as possible, they should be related to more specific units of performance. These are operational and maintenance tasks (or HPF's in HRTES). Therefore (in Chapter 8--Diagnosis) HRTES presents highly detailed methodologies for answering the question--How effective was each human resource area (training, HFE, and personnel selection) for each significant task which was performed inadequately? Using this approach and answering this much more specific question will produce much more useful results than answering the question--How effective was each human resource area for overall system performance? If you fully use Chapter 8, this more specific question will be answered.

However if you must answer the general issues--How effective were: training, HFE, and personnel selection for the system as a whole?--then two further techniques are available. Both techniques depend upon the use of Chapter 8.

When Chapter 8 has been completed you will have listed those HPF's which were performed inadequately and which, according to an analysis of the Evaluation Tree, were of high significance for the system. For each of these HPF's you will have a series of hierarchically nested Indeces of

Adequacy for each of the three human resource areas (Training, HFE, and Personnel Selection). These nested indeces will range from small, specific segments of a given human resource area, as applied to that HPF, to the whole human resource area as applied to that HPF. Each Index of Adequacy will range from 0-100, 100 meaning total adequacy of that human resource area, or segment, for that HPF.

Both techniques that will be described for specifying overall effectiveness of a human resource area use the Indences of Adequacy, but they differ according to the relative weight given to each Index.

Technique 1

In the first technique one must make two assumptions: (1) All HPF's have the same importance for the system and therefore all human resource area Indeces of Adequacy have the same importance. (2) If an HPF was successfully performed, its human resource areas are fully adequate by definition, and if it was unsuccessfully performed (but not important enough to diagnose) one does not care about its human resource areas. This latter assumption is necessitated by the absence of data about human resource areas of HPF's which were not diagnosed. Hopefully there will be few or no HPF's in the category of--not important enough to diagnose, but inadequately performed.

(1) In this technique for each human resource area you assign 100, to each HPF that was performed adequately.

(2) You eliminate all HPF's that were performed inadequately, but which were not important enough to diagnose.

(3) You then compute the mean of the remaining HPF's (using 100's for adequate HPF's and Indeses of Adequacy for those HPF's diagnosed.

If this mean were 100 that human resource area' effectiveness could be considered fully adequate. The farther that the mean fell below 100 the less effective the human resource area for the system as a whole.

Technique 2

The second technique also depends upon two assumptions. The first assumption is that HPF's have differing importance for the system and therefore all

human resource area Indeces of Adequacy have differing importance depending upon the importance of the task, SPI, mission, etc. The second assumption is identical to the second assumption described above, since it is necessitated by the same lack of human resouce area data for any but inadequately performed, significant HPF's.

(1) For a given human resource area you first eliminate any HPF's from the Evaluation Tree that were performed below criterion, but that were not significant enough to be diagnosed in Chapter 8.

(2) Next, if any such HPF's were eliminated, you must recompute the evaluation weights of the remaining HPF's in the families of those HPF's eliminated. As you remember all HPF's have identical weights, and those in the same family have weights which sum to one.

(3) You then assign a 100 value, by definition, to all HPF's which were performed adequately.

(4) Finally, using the 100's, and the actual Indeces of Adequacy for those HPF's which were diagnosed, you fold back the Evaluation Tree according to the instructions on pages H7-17 and H7-18, just as you did for the "P Values." If this procedure is followed for each of the three human resource areas, it will result in estimates of adequacy of overall training, HFE, and personnel selection as applied to: each HPF, each HPF-Group, each SPI, each System Function, and the system in general. For each node completely adequate treatment of a human resource area would result in a an index of <u>100</u>, not 50 as in the computation of performance values. The farther below 100 an index at a given node, the less adequate that human resource area's treatment at that node. Thus general human resource effectiveness issues can be dealt with in HRTES.

Conclusion

As a general rule, it is preferable to use neither of these two techniques, and to use Chapter 8 as written. That is, it is preferable to determine effectiveness of each human resource area for each inadequately performed and significant HPF. If equivalently detailed human resource data were taken for all HPF's performed, then these techniques would be recommended. However as this is a very time consuming process it will probably not be done. The result is the necessity for assuming adequate human resource treatment of HPF's that were performed above criterion, plus the elimination of HPF's that failed and were not significant. Such an assumption is operationally useful, but its validity is questionable. However if you must produce such a general human resource estimate, and if you cannot take the time to gather human resource data of high quality on all HPF's, these two techniques may serve. You will have to decide which of these two most nearly fits your model of appropriate human resource evaluation. In general the second technique (using the existing Evaluation Tree) is recommended on the ground that assigning more adequate human resource area treatment to more significant HPF's should result in more effective overall assessment of that human resource area for the system as a whole.

8. DIAGNOSIS

CONTENTS:

I. Determining Which HPF's Should Be Diagnosed.

II. Determining Which Diagnostic Measures to Apply to HPF's.

III. Diagnosing the Causes of Subcriterion HPF Performance.

ACTIONS:

(1) Deciding which Evaluation Tree nodes to diagnose.

(2) Tracing of selected nodes to HPF Nodes.

(3) Determining hardware or human resource area cause(s) of HPF performance.

(4) Deciding strategy for taking human resource area measures.

(5) Determining which class of diagnosis measures to apply to HPF's--Expert Measures or Questionnaire Measures.

(6) Gathering Worksheets and other documentation for diagnosis.

(7) Selecting experts in each of the three diagnosis areas: training, HFE, and personnel selection (if this class of measures is selected).

(8) Copying appropriate blocks of diagnostic measures and sending them to selected experts along with supporting Worksheets and documentation.

(9) Diagnosing the cause(s) of inadequate performance of each selected HPF based on material returned from experts, or Questionnaire Measures.

(10) Writing of diagnosis for each selected HPF.

(11) Summarization of critical incident data.

PRODUCTS:

(1) Indices of Adequacy of each human resource area and component indices of Adequacy for each selected HPF.

(2) Diagnosis of the probable cause(s) of subcriterion performance of significant HPF's.

(3) Summarized critical incident data.

USED AS: (1) Working papers for preparation of Independent Evaluation Report.

I. DETERMINING WHICH HPF'S SHOULD BE DIAGNOSED

DESCRIPTION: Performance values of HPF's are the basis of performance values of all the nodes of the Evaluation Tree. Therefore, any upper level node of the Tree which was performed inadequately actually must be diagnosed at the HPF level. In this section you will determine if any performance is significant enough and inadequate enough to require diagnosis. If this is the case, you will determine which HPF(s) was responsible for the inadequate performance.

PROCEDURE:

(1) Examine the completed Evaluation Tree.

(2) Identify and mark those nodes of the Tree that have a performance value (P) and an upper confidence limit value (U) less than 50.

(3) Start at the highest level of the Tree that has marked nodes and decide if each marked node is significant enough to deserve diagnosis. This decision should be based on your knowledge of the system that was tested. Also, see the Detailed Explanation.

(4) Once you have selected marked, higher level nodes to be diagnosed, trace your way down through the branching structure to the next marked node.

(5) Continue the tracking process until you reach the marked HPF(s). You have now traced your way down through the Tree's branching structure from a marked node that is significant enough to be diagnosed to the marked HPF node(s). It is this marked HPF node(s) that caused the higher level node to receive a performance value less than 50.

DETAILED EXPLANATION: For a detailed explanation of: marking inadequate nodes, determining the significance of nodes, and tracing through the Evaluation Tree, see pages H9-61 and H9-62.

II. DETERMINING WHICH DIAGNOSIS MEASURES TO APPLY TO HPF'S

DESCRIPTION: This chapter contains two types of measures:

(1) expert measures which are taken by training and Human Factors Engineering personnel and are objective for the most part; and

(2) questionnaire measures which are the scales found on the "Opinion Summary Data Worksheets" for the HPF's being diagnosed and which are based on the opinions of field test players and observers. Expert measures are considerably more desirable although they are more time consuming than questionnaire measures.

Three human resource areas are measured in this chapter: training, human factors engineering, and personnel selection. There are expert measures of each of these three human resource areas. There are questionnaire measures of training, and of the human-machine interface (HFE or Personnel Characteristics). It is not useful to ask someone if they had difficulty operating a device due to their lack of ability. Therefore, questionnaire measures of the human-machine interface cannot differentiate between causes of inadequate HFE design and inadequate personnel characteristics. They can only point to the specific problem, not its cause.

You now have to decide whether:

(1) to perform the diagnosis entirely with questionnaire measures;
(2) questionnaire measure scores can be used to reduce the number of expert measures that your training and HFE experts will take;
(3) training and HFE experts will have to decide which expert measures to take without the aid of questionnaire scores;
(4) personnel selection measures are applicable to this situation.

PROCEDURE:

(1) To get an idea of the material covered by the expert measures, look through the "Expert Measure Index" on pages H8-6 through H8-8.

(2) If more than one HPF is to be diagnosed, and you are going to use expert measures, perform the following subprocedure:

(a) If your expert(s) already has a copy of HRTES, send copies of the: appropriate "Summary Data Worksheets" with scores below 50 circled, an "HPF Diagnostic Worksheet (page W8-4) with HPF's, etc., filled in, and a statement of which block of diagnostic measures (training, HPF, or Personnel Selection) are to be considered.

(b) If your expert(s) does not have a copy of HRTES, first make <u>one</u> copy of the block of diagnostic measures, no matter how many HPF's are to be diagnosed. Then make completed copies of the "Summary Data Worksheets" and "HPF Diagnostic Worksheets" as above. Make sure you include copies with each block of diagnostic measures sent to experts.

(3) If expert diagnostic measures of training are to be considered, you will have to include copies of the following in your submissions to the training expert:

(a) the completed "Evaluation Tree;"

(b) the set of completed "OT Training Data Collection Worksheets" that were filled in by OT trainers following OT training.

(c) "HPF Difficulty Worksheets" completed for the HPF's to be diagnosed.

EXPERT DIAGNOSTIC MEASURE INDEX

[1] Two alternative expert measures are available.

[2] This is a figure of merit that is based on preceeding diagnostic measures. It is to be computed by diagnostic measure experts following measure completion.

EXPERT DIAGNOSTIC MEASURE INDEX (CONT.)

EXPERT DIAGNOSTIC MEASURE INDEX (CONT.)

(4) If you have decided to use questionnaire measures for diagnosis, you will have to retrieve the "Opinion Summary Data Worksheets" for those HPF's. The individual scale scores are the questionnaire measure outcomes for each HPF being diagnosed. More specific information can be obtained by examining the actual questionnaires which were completed.

DETAILED EXPLANATION: For a detailed explanation of diagnostic measures see pages H9-66 through H9-68 and Chapter W8 in the Workbook.

IV. DIAGNOSING THE CAUSES OF SUBCRITERION HPF PERFORMANCE

You now have received the three completed Summary Worksheets for Diagnosis (Training, HFE, and Personnel Selection), or that subset of the three which was felt to be appropriate to the HPF being diagnosed, plus all the measure worksheets upon which they were based. Each completed Diagnosis Worksheet contains a *single Index of Adequacy* for its human resource area (as applied to the HPF), the individual, specific component Indeces of Adequacy, and the specific problem areas which caused component Indeces to be less than 100. It should be remembered that an index of 100 indicates complete adequacy, not optimum design or personnel characteristics. These completed Diagnosis Worksheets plus any Critical Incident Worksheets (for incidents which occurred at the time of HPF performance) can now be used for diagnosing the causes of the subcriterion, significant HPF performance.

The type of diagnosis that results from this procedure is dependent upon your goal, and therefore to some extent upon the audience for whom it is being prepared. The specific problem areas, plus any relevant critical incident information, will tell you all the negative aspects of the three human resource areas which contributed to the inadequate HPF performance.

However, it is the various collective Indeces of Adequacy which will aid you in making statements about overall adequacy of larger units. That is, much totally adequate design, in a given human resource area, to some extent, makes up for a small amount of inadequate design in a low criticality segment of that area. Both criticality and extent of adequacy of individual segments is taken into account in the various collective indeces. Using the statements of specific problem areas (plus their criticality weights when available), the various intermediate collective indeces and the indeces for the human resource areas, it should be possible for you to include the following information in your diagnosis of a given inadequately performed HPF.

(1) Specific inadequacies in training, HFE, and personnel characteristics which led to the subcriterion performance of the HPF (in most cases each listed inadequacy would include the criticality of the inadequate human resource element).

(2) Critical incidents which took place at the time of HPF performance which may have led to, or played a role in, the inadequate performance.

(3) Level of adequacy of intermediate, functionally related human resource elements, taken together. Such a level of adequacy would include criticality and adequacy of the component human resource elements in these functionally related groupings. Such intermediate groupings could include: training time allocation, training methods adequacy, display readability, display usability as a whole, control accessibility, control usability as a whole, aptitude, cognition ability as a whole, size for accessibility, control accessibility, and many others. Such

intermediate indeces exist, in HRTES, at various hierarchical levels, but any index between actual problem area and human resource level index is considered intermediate. Since these intermediate indeces would all be on the same 0-100 scale of adequacy, comparisons would be possible, across human resource areas. These intermediate indeces would allow well designed human resource elements to make up for inadequately designed ones, and would make the relative criticality of these elements a major factor in resulting overall balance of adequacy.

(4) Level of adequacy of each of the three human resource areas in relation to the inadequately performed HPF. These indeces of adequacy would have the same properties as the functionally related, intermediate indeces, of which they are composed. They would indicate the level of adequacy of training, HFE, and personnel selection for the HPF being diagnosed. They would allow well designed elements to make up for inadequately designed ones, to some extent, and would use criticality of elements as the major factor in determining to what extent good design could make up for inadequate design.

With this information you should be able to diagnose the specific causes of inadequate HPF performance, and to describe the adequacy of functionally related design areas. It is suggested that you include the actual Diagnosis Worksheets as part of your analysis. In addition, it is suggested that you retain copies of both the Diagnosis Worksheets and the various human resource measure worksheets, which were completed, as part of the permanent record of this test.

9. APPENDIX

CONTENTS:

ACTIONS:

(1) Applying questions about the development and status of Human Factors Engineer to the system to be tested and obtaining answers from appropriate individuals.

(2) Answering questions and taking measures about projected training for the system to be tested.

(3) Comparing the projected full-scale training package with the OT training package.

(4) Answering questions about projected OT training to encourage the performance of the appropriate developmental procedures, and to determine if adequate training will be available by the start of the field test.

(5) Answering questions about manpower planning and selection to encourage the performance of the appropriate developmental procedures, and to determine if appropriate participants for the field test will be available on time.

(6) Gaining detailed understanding of problems of diagnosis.

PRODUCTS: Products in the Appendix will be a series of various questions with specified answers.

USED FOR: Used for monitoring the ongoing development of the various elements of the system to insure that the important elements are actually performed. Used to gain detailed understanding of diagnosis procedures.

I. TIMELY INCORPORATION OF HUMAN FACTORS ENGINEERING INTO SYSTEM DESIGN

From the point of view of the operator or maintainer, the Human Factors Engineering (HFE) of a System is the System. That is, the part of a System with which the operator/maintainer interacts is the part which is of greatest concern to him. This section is designed to aid the user in insuring that HFE requirements are being incorporated into the System sufficiently early enough to impact significantly the ultimate design features, leading to a more meaningful Operational Test. The questions are adapted from Holshouser (1975). In most cases, you will not be able to asnwer these questions yourself. You will have to obtain the answers from either appropriate members of government HFE facilities of individuals who are part of the organization that is designing the System. This may not be particularly easy, but the simple fact of your attempting to obtain these answers at an early stage of System development will increase the likelihood that they will be carried out properly, thus permitting positive answers.

HUMAN FACTORS ENGINEERING EVALUATION QUESTIONS

	YES	NO	IN PROCESS
1. Will the anticipated working environment, including the physical aspects (weather, illumination, temperature, humidity, ventilation, noise, vibration, ionizing radiation, etc.) and the operational aspects (high density of threat, operational communications, work loads, duty cycles, etc.) adversely affect operator performances?			
(a) What types of effects are to be expected (reduced visual/ auditory field, reduced tracking ability, or reduced joint mobility?			

	YES	NO	IN PROCESS
(b) If the effects are critical to mission performance, how can these expected reductions in performance be minimized?			
2. Are tests identified which will determine how well the human operator has been integrated with the System elements?			
(a) Will tests permit determination of whether operational requirements have been met?			
3. Have alternate programs been identified which might interfere with the one under development? If so, what studies are necessary to determine whether those programs are compatible with each other in terms of signal processing, information flow,			

	YES	NO	IN PROCESS
feedback lines, lines of authority or control and ILS (integrated logistics support) requirements?			
(a) Have human factor problems (if any) encountered in the deployment of similar Systems currently in use been identified?			
(b) Have the human interfaces between programs been determined?			
4. Are man-machine interfaces defined and areas critical to success of System mission pointed out? Are trade-off studies (such as alternate allocation of function schemes or alternate hardware/software designs)			

	YES	NO	IN PROCESS
for man-machine interface interfaces in the System planned and discussed?			
5. What testing technique/procedure will most efficiently answer the critical HFE questions and issues (laboratory testing, part-task simulations, flight testing, etc.)?			
Have analyses and studies been accomplished on the equipment design to determine whether the equipment characteristics demand operator performance which exceeds human capabilities or approaches limitation which may significantly contribute to the occurrence of one or more of the following conditions (but is not necessarily limited thereto)?			

	YES	NO	IN PROCESS
(a) Jeopardize mission performance.			
(b) Degradation of System accuracy.			
(c) Delay beyond acceptable time limits.			
(d) Improper operation leading to System failure.			
(e) Results in excessive maintenance and down time.			
(f) Degradation below reliability requirements.			
(g) Damage to equipment.			
(h) Compromise of System security.			

	YES	NO	IN PROCESS
(i) Injury to personnel.			
Have back-up modes or compensatory aids been considered in the event of partial System failure?			
6. Have the test plans been detailed sufficiently to describe the conditions of test, control of variables, data collection techniques, and method of analysis of results?			
(a) Will the conduct of tests produce results which will identify deficiencies, difficulties, limitations, and shortcomings?			
(b) What are the shortcomings of the proposed tests?			

	YES	NO	IN PROCESS
7. Have the critical issues and questions regarding the impact of the human operator on System operations been addressed in terms of manning levels, skill levels, workloads, duty cycles, stress, and extremes of environment? Have back-up modes or compensatory aids also been addressed?			
8. Has the developing agency specified a planned schedule of events with sufficient detail to plan the HFE test time table? Are the milestones to be met through testing attainable within time and money constraints? If not, what alternate plans are being considered?			
9. What limits have been established for the System in respect to human performance; i.e., detection ranges, lock-on ranges, response or reaction times, update times, etc.?			

	YES	NO	IN PROCESS
(a) Are these limits within the range of capability of the human operator?			
(b) What are the probabilities of System failure?			
(c) What back-up modes or compensatory aids are planned in the event of partial System failure?			
10. Have the various disciplinary teams, including Human Factors Engineering, provided the Program Coordinator with sufficient information from their R&D			

	YES	NO	IN PROCESS
tests, investigations and demonstrations on breadboard and prototype models for him to establish a working relationship with other groups who will be involved in the System development? Has the impact of new and/or unique System items on the human operator been determined?			
11. Has preliminary or research testing considered or identified potential HFE areas where additional emphasis could result in improved System performance? If so, what are the results?			
Have the human factors R&D investigations surveyed the state-of-the-art in control/display Systems, and other man-machine relationships?			

	YES	NO	IN PROCESS
(a) Have the investigations considered research which advances the state-of-the-art?			
(b) What efforts have been made to determine the probability of success and the impact on System delivery time for those items which advance the state-of-the-art?			
12. Has an analysis been made of techniques which could be used to degrade the information available to the System operator or otherwise render him ineffective; i.e., generating false targets?			
(a) What consideration has been made to provide alternate modes of operation?			

	YES	NO	IN PROCESS
(b) What consideration has been given to aid operators in recognizing that countermeasures are being used against them?			
13. What trade-offs were considered in the allocation of automatic versus manual functions in respect to counter-countermeasures?			
(a) If the System is an "add-on," "off the shelf," GFE (government-			

	YES	NO	IN PROCESS
furnished equipment) or CFE (contractor-furnished equipment) System, what demonstrations were conducted to determine that the human operator is adequately integrated into the primary System as well as the add-on System?			
14. Have the psycho-physical effects of *electromagnetic* radiation and interference been assessed? Have compatibility tests been planned to determine the presence of interference from on-board equipment which effect the operator's displays?			
(a) Have test criteria been established in respect to signal-to-noise ratios and detection thresholds?			

	YES	NO	IN PROCESS
Have signal intensities been established for the operator in a tactical operating environment?			
15. Have the human performance estimates of critical functions been validated to assure that no adverse affects will occur in terms of cost, reliability, efficiency, effectiveness, and safety?			
(a) Have investigations or tests verified that previous requirements have been satisfied and/or have not changed significantly to affect service suitability?			
16. Have the limits of the System in respect to human performance been verified; i.e., turn-around times, reacquisition time, loading time (computer and weapon), etc?			

	YES	NO	IN PROCESS
(a) Have demonstrations or test data shown that these limits are acceptable for both the human operator and the System?			
(b) Have test results indicated that the task loading is acceptable for the human operator?			
17. Have analyses, studies, tests, experiments, and/or demonstrations been performed to provide data on the effective information flow and processing, including decisions and actions, required to accomplish the System objective?			
(a) Have analyses and trade-off studies been conducted to determine which System Functions should be machine-implemented and which should be assigned to the human operator/maintainer?			

	YES	NO	IN PROCESS
(b) Do these analyses indicate that the specific requirements can be met?			
18. Have sufficient numbers of representative subjects weapon Systems, targets, environmental conditions, tactical situations, and combinations thereof been employed to provide data to ensure a valid, overall evaluation of total System performance?			
(a) Was the sample size adequate?			
(b) Were appropriate test techniques used?			
19. Are data, including human error data, provided to determine degradation of System operation below reliability requirements?			

	YES	NO	IN PROCESS
(a) Are the test data adequate for back-up of contract requirements for reliability?			
(b) Are failure analyses performed or failure data collected to differentiate between failures due to equipment alone, man-equipment incompatibilities, and those due to human error?			
20. What new problems have been identified as a result of the developmental tests? What is the significance of the problem to the full scale development decision? Have previous studies, analysis, and testing of man-machine interfaces provided sufficient information to select the best arrangement in keeping with program costs and direction?			

	YES	NO	IN PROCESS
(a) Have workload levels been determined, and are they acceptable?			
(b) Has information processing been studied; i.e., type and quantity of information. Are the processing requirements within the operator's capability?			
21. What determinations have been made to assure that the operator/maintainer is not being exposed to the danger of fire, explosion, fungus, toxics, or other debilitating effects from new materials?			
22. Where critical materials are employed in man-equipment interfaces, have the advantages and disadvantages of that material over conventional material been assessed in terms of			

	YES	NO	IN PROCESS
procurement time and cost, and the impact on operation and maintenance?			
23. Has a survey of the state-of-the-art been conducted to identify new and innovative man-machine interface devices and techniques? If so, have trade-off studies been performed to determine their applicability to the weapon System?			
24. Have all interfaces been described in sufficient detail to plan compatibility tests of inter-intra interference; visual, auditory, and physical access, etc? Is the contractor furnished equipment which provides the hardware interface with the operator/maintainer qualified or otherwise certified? If not qualified, what studies should be performed to either qualify it or quantify the effect on operator/maintainer performance, costs and delivery time? Have the human factors plans been			

	YES	NO	IN PROCESS
detailed for incorporation into the operational test plans for the System integration tests?			
25. Have the human factor aspects of logistics such as maintenance, training, manuals, and personnel been defined and incorporated into the ILS program plan? How do these aspects affect the deployment date and life cycle costs?			
26. Have the critical human factors questions and issues to be resolved prior to Milestone III been refined and adequately addressed? What additional questions or issues need to be before the production decision?			

	YES	NO	IN PROCESS
27. Are there any parts of the subsystem under development which should be appraised by human factors engineering for potential impact on user performance? Are there any substitutes which should be tested as an alternate?			
28. Has the data package documented the type of testing, test results, and conclusions in terms of ease of unambiguous operation, high reliability, ease of check out, removal, and replacement? Is the documentation comprehensive enough to support the			

	YES	NO	IN PROCESS
decision to accept or reject the System?			
29. Have the human factor tests of user acceptability or operability been integrated with the reliability and maintainability? Have these tests been performed in a simulated operational environment utilizing service personnel?			
30. Have all the HFE tests provided sufficient data for resolving the critical issues? Have new problems been identified as a result of the developmental tests? If so, what is the significance of the problem to System production?			

---

31. Have fixes to human factors problems, identified during developmental			

	YES	NO	IN PROCESS
testing, been validated by further testing in a simulated operational environment?			
32. Have human factors compatibility tests covered the integration with the other Systems? Have the tests been sufficiently detailed and realistic to cover the interfaces within and between subsystems and Systems?			
33. Has analysis of the developmental tests data or Operational Test data been performed to determine the status of critical areas identified in the advanced development and engineering model? Have studies been conducted to determine the correlation of test data obtained from an engineering development model to the final production model?			
34. Were the human factors tests designed to duplicate or simulate the anticipated operational environment? Were tests designed to maxi-			

	YES	NO	IN PROCESS
mize the usefulness of the data by manipulation of System function variables? What differences exist between the environment in which the HFE tests were performed and the expected operational environment? What impact will the differences have on the operational use of the System?			
35. Have the Operational Tests demonstrated that the System can be effectively operated and maintained by the level of personnel skills, manning levels, workloads, and duty cycles anticipated to be available under service conditions? Have the demonstrations affected the training and plan maintenance concepts?			
36. Have tests been conducted in unfamiliar territory against unfamiliar			

	YES	NO	IN PROCESS
and uncooperating targets? Have the tests demonstrated that the operator/ maintainer can perform the required task on his own without help from outside advisors or at least with only that support which would be available in the operational environment?			
37. Have the planned tests been designed to determine the extent of degradation which will occur in the operational environment? Do the test results concur with O&M cost estimates? What tests are planned to check differences between prototype and production models and will the OT&E tests cover ECPs (engineering change proposals) and program change orders? Have the planned tests specified the manning levels and skill levels of the service personnel who will use and maintain the System?			

	YES	NO	IN PROCESS
38. Is all threat information within the crew's sensing and processing abilities, and are they readable under all lighting and tactical environments? Are appropriate warning devices available to indicate the use of countermeasures?			
39. Has the System been tested against unfamiliar targets in unfamiliar territory? That is, are the targets representative, and are the test operations representative of simulated combat in respect to target density and activity? Are unique briefing materials required to maximize System utilization?			
40. Have tests been conducted which compare the new operator/maintainer System interface with the operator/maintainer-System interface of the System being replaced?			
(a) Are there tests which were conducted on the engineering, development, prototype, pilot or limited production models which need to be replicated on			

	YES	NO	IN PROCESS
the full scale production model because of significant changes which occurred during the evaluation of the new System?			
41. Have all modes of operation been tested which require operator intervention? Were the operator assisted modes tested under realistic loads, stress, and environmental conditions? Did the results of these tests have an impact on the production decision?			
42. Have specific support equipment, test, gear and techniques been designed to inform the operator/maintainer of System life status in a direct unambiguous and nondestructive manner? Have level of degradation criteria been established and promulgated? Has a policy been provided regarding corrective action?			
43. Has the ancillary equipment interface with the System been tested for compatibility and accessibility of cables, controls, displays, power			

	YES	NO	IN PROCESS
sources, and the environmental effects of noise, light, vibration, motion, etc?			
44. Were the appropriate criteria specified in relation to the operator/ maintainer environment in a System safety program?			

II. Effect of System Design on Training.

To insure that OT training is a realistic sample of the full scale training package, the nature of the package must first have been defined. This section is constructed to monitor the impact of the System's design on the full scale training package design. This is to insure that there will at least be a fully worked out training concept which can be abstracted for OT training. Read the questions, and obtain answers to the measures connected to the questions from appropriate training experts.

1. Rationale. A major element in the life cycle cost of a System is the cost of developing and carrying out the training program associated with the System.

The design of the System has a direct impact on the training requirements for that System. A System which includes sophisticated computers to aid operators, for example, may require additional operator training to learn to operate the computer, or it may simplify the training by making the task easier. A System that is similar in operation to

existing Systems may facilitate the development of the training package by capitalizing on existing material.

Two competing versions of a System may be equal in terms of their operational performance yet result in vastly different life cycle training costs. The overall cost-benefit analysis of a System must, therefore, address training. Training, however, is not independent of design. The earlier in the design cycle the issue of training implications is addressed, the easier and less expensive it is to modify the System to minimize life cycle costs. Often small design changes can have significant impacts on training. A change in the aiming device of a weapon might reduce training time by several hours and live fire practice by a significant percentage.

2. Approach. This section of HRTES is not intended to set forth a methodology for assessing the impact of System design on training. It is not always easy to identify those aspects of a particular System which, if modified, could result in training cost savings. This often involves the judgment of persons familiar with military training on Systems similar to the one under development. Even assessing the impact of a given design on the various aspects of a training program is, to a great extent, an art. The purpose of this section of HRTES, therefore, is to get you to ask the question early in the design process so that tradeoffs may be effected. Although the methodology for answering the questions is not prescribed, a list of critical training dimensions which can be affected by System design is given. This will allow you to evaluate whether all relevant aspects of training have been addressed in the assessment of the impact of System design.

3. Questions. Two questions are given here: the first one is addressed more to TRADOC than to the System designer, while the second is

more a question for the System designer than TRADOC. Bear in mind, however, that the answer to either question would require dialog between TRADOC and the System designer if valid answers are to be obtained.

QUESTION 1: What is the impact of the System design on the training function?

QUESTION 2: To what extent have alternative System designs been evaluated in terms of their impact on training?

4. Measures. This list of measures includes the principal parameters of training that may be impacted by the particular design of the System. Given a specific design, or alternative designs, TRADOC personnel should be able to assess the impact of each on these parameters. In most cases, the assessment will be subjective, based on prior training experience with similar Systems. Some of the measures involve quantitative measurement (e.g., calendar time, number of devices), while others are more qualitative in nature (e.g., level of fidelity, real estate requirements).

Ultimately, all of these measures could be reduced to dollar cost figures for each aspect of training.

(1) Training Time: This aspect of training addresses the total amount of time required to train an operator to an adequate level of proficiency. This includes all forms of training (e.g., school, unit, on-the-job).

(a) Minimum number of training hours required to train to criteria.

(b) Minimum total calendar time required to train an individual or team to criteria.

This measure is somewhat different from the first measure. It may take 40 hours to train, but that 40 hours could be done in one week or spread over one month. This measure has importance in determining throughput availability of manpower to operate the fielded System.

(2) Training Device Requirements: Training devices include simulators, mock-ups, and actual Systems used in training. In complex Systems, this category can involve considerable capital outlay.

(a) Minimum time required on each type of training device.

(b) Minimum level of fidelity required to maximize transfer for type of device.

(c) Minimum number of each type of device needed to maintain an adequate output of trained personnel.

(d) Total cost for training devices.

This is really a combination of the previous three measures into a single metric. The other measures will be valuable in pinpointing the cause of the costs and should be used if possible.

(3) Training Materials: Training materials include those things used by the instructor or student that would not be classified as devices. This includes audio-visual materials, books, tests, and other printed materials.

(a) Minimum amount and type of materials required to carry out training.

(b) Minimum amount of person-hours required to produce the materials.

(c) Total cost of materials.

(4) Real Estate Requirements: This refers to buildings, classrooms, etc., needed to carry out training and geographic requirements necessary for skill practice. Such things as firing ranges with specific terrains would be included.

(a) Geographic size and location requirements of facilities.

(b) Space and facility requirements, (e.g., square feet of space required, special requirements for buildings on terrain features).

(c) Cost of real estate and facilities.

(5) Trainer Characteristics: Trainers must be trained and they must possess specific skill levels. A System may be designed that reduces the training and skill requirement,

thereby effecting a cost savings and insuring an available supply of trainers.

(a) Minimum number of trainers required to maintain an adequate supply of trained operators.

(b) Minimum amount of time required to train the trainers.

(c) Minimum skill level of trainers as a trainer, (e.g., number of hours of experience as a trainer required).

(d) Minimum skill level of trainer on the generic System class. This would include number of hours of System operator required; MOS skill level required.

(6) Organization of Training: This category deals with the organization of the training program.

(a) Minimum throughput required. (This is measured in terms of number of people to be trained per unit of calendar time.)

(b) Maximum student-instructor ratio.

(c) Proportion of time to be spent in various types of training. Including entry school, special school, OJT including embedded training, individual or group training, unit training, team training, engagement training.)

(d) Time to be spent in various instructional modes. (Including classroom, self-paced, simulation, actual equipment.)

(7) Evaluation Requirements: This addresses the extent of performance evaluation necessary to determine adequacy of the trainees.

(a) Type of evaluation mode needed. (Includes instructor ratings, objective paper, pencil tests, training device objective measures.)

(b) Minimum time devoted to evaluation activities.

III. Similarity of OT Training to Full Scale Training.

To determine if a new piece of hardware will be effective in the field, OT must use a training package which closely resembles the projected full-scale training package. To do this, it is necessary that there be a reasonable idea of the nature of the full-scale training package in advance of the OT. The more closely the training of the OT operator/maintainer population resembles that for the fielded System, the greater the validity of the data taken in the OT, and the more useful the resultant evaluation. Unless one can state with some confidence that the OT training package is a reasonable representation of the full-scale package, one should seriously consider postponing the OT until a valid sample can be provided.

To answer the question of the resemblance of OT training and full-scale training, one must first be able to give an affirmative answer to the following question: Will enough of the full-scale training package be available to permit the OT training package to be designed to include its most significant characteristics? If the full-scale training package is not available, however, a detailed training design outline which specifies the full-scale characteristics may suffice.

Data for the following parameters of training should be obtained from appropriate training experts. The following are hte parameters that should be considered when comparing the OT and full-scale training packages. The full-scale training package should be analyzed according to these parameters. The OT training package should then be designed so as to resemble the full-scale training package along these parameters. The greater the differences between the two packages, the less likely the OT will produce valid results.

Training Package Parameters

1. Time

1.1 Total training time (in training hours).

1.2 Total training time devoted to each critical task.

1.3 For each critical HPF, the percentage of time devoted to:

1.3.1 Classroom.

1.3.2 Training devices/simulators.

1.3.3 Field Systems.

1.3.4 Self-paced teaching machines.

1.3.5 Team training.

NOTE: It is possible for 1.3.1-1.3.5 to sum to more than 100%.

2. Critical HPF's

2.1 Critical HPF's to be trained.

NOTE: If necessary, simulators may be replaced by field Systems. Parameters should reflect this replacement.

2.2 Student-trainer ratio for each critical task.

3. Conditions

3.1 Conditions in which training will be performed.

3.2 Critical HPF's to be performed in each condition.

4. Trainers

4.1 Trainers' MOS skill level.

4.2 Trainers' training exerience (in months).

4.3 Trainers' experience in this or similar field Systems as an operator/maintainer (in either hours or months, as appropriate).

5. Student Evaluation

5.1 The method for determining student proficiency on each critical HPF.

5.2 Exit criterion performance required for each critical HPF.

IV. Timely Development of OT Training.

The training of operators and maintainers is of significant importance in the OT process. The development of the OT training package is a series of actions, many of which are dependent on each other for completion. In order that OT training be ready in time for the OT exercise, careful monitoring of the training development process must be performed from a fairly early stage of development. This section presents you with a series of questions about the development of the OT training package which should be answered as the process of comparing OT Training to Full Scale Training continues.

1. Development of OT Training. This section deals with two questions critical to the development of a training package for OT:

	YES	NO	IN PROCESS
1. Will a training package be ready on time for conducting the OT?			
2. To what extent is the OT training package a reasonable representation of the projected full-scale package?			

Each of these questions will be addressed separately in this section. In reality, however, they must be considered together to insure that a representative package is developed in time for the actual operational test.

The design of System hardware is only one aspect of operational effectiveness. Hardware does not operate in a vacuum. It is operated and maintained by personnel. The training of these personnel contributes directly to the operational effectiveness of the System. The purpose of OT&E is the evaluation of System effectiveness in the anticipated operational environment, although it is not the function of OT to evaluate hardware effectiveness when operated or maintained by a population that is significantly different from that which it will deal with in the anticipated operational environment.

It is critical, therefore, that the OT training package be prepared with the same degree of diligence as the full-scale package will be. OT training cannot be taken lightly. A training package hastily developed for OT may result in an invalid test.

2. Timely Development of an OT Training Package.

	YES	NO	IN PROCESS
1. Has a schedule for the development and preparation of the OT training package been produced?			
2. Have adequate task analyses been completed for all appropriate personnel to guide in the development of the OT training package?			

	YES	NO	IN PROCESS
(a) Have performance criteria been determined for the task elements identified?			
(b) Have tasks been assigned criticality and difficulty ratings?			
(c) Have probable and worst case operating conditions been identified?			
3. Has a curriculum plan been developed?			
(a) Has training time been allocated among the various types of training (e.g., classroom, individual, self-paced, simulation, operational systems, and/or team training)?			
(b) Have tasks been allocated among various types of training?			

	YES	NO	IN PROCESS
4. Will appropriate training devices be available to train OT participants?			
(a) Have training device characteristics been matched to the tasks and conditions selected for OT?			
(b) Can existing training devices from similar Systems be used in OT training with a minimum of alteration?			
(c) Can the new System be used as a training device for OT training?			
(d) Will there be a sufficient number of training devices/new Systems available to train all OT participants?			
5. Will appropriate training materials (audio-visual/printed matter) be available to train OT participants?			

	YES	NO	IN PROCESS
(a) Can training materials from existing similar Systems be used in OT training with a minimum of alteration?			
(b) Will there be a sufficient quantity of training materials available to train all OT participants?			
6. Will there be a sufficient number of trainers available to carry out OT training?			
(a) Has a potential group of trainers been identified with the appropriate characteristics that will be available to carry out OT training?			
(b) Can the trainees be trained in time for OT training?			

	YES	NO	IN PROCESS
7. Will there be a sufficient number of appropriate individuals available to be trained for OT?			
8. Have resources been allocated to insure the timely development of the OT training package?			
(a) Have OT training resources been allocated to reflect the relative criticality and difficulty of the tasks and conditions?			
(b) Have sufficient resources been allocated to permit adequate operational practice (including live firings) during OT training?			
(c) Has the full cost of the OT training package been projected and sufficient funds allocated to execute it?			

V. Effect of System Design on Manpower Planning and Selection.

This section is designed to aid you in determining that the manpower requirements are being incorporated into the System sufficiently early to impact the design features. The answers to the questions in this section will tell you that manpower planning is being conducted correctly and in a timely manner, leading to a meaningful OT. To determine that the OT participants consist of individuals with a mix of skills representing the real user population, the skills required of this user population must first have been determined. The flowchart found on page W8-43 of the Workbook aids in this process. Answering of the questions and the working of the flowchart should be done by appropriate experts.

WHAT ARE THE REQUIRED HUMAN ABILITIES AND SKILLS

	YES	NO	IN PROCESS
1. Have task analyses been accomplished in order to determine the specific human functions required for effective System performance?			
2. If task analytic data are not available or obtainable, are there task analytic data from similar Systems or information from experts which can be used to specify ability/skill requirements?			
3. Are the data derived from task analyses or other sources adequate, reliable, comprehensive, recent, etc.?			
4. Have specific human performance criteria been established for each			

	YES	NO	IN PROCESS
task such that failure to meet these criteria would degrade System performance?			
5. Have the tasks been classified or structured to generate clusters having common elements in order to simplify skill assessment?			
6. What human abilities or skills are necessary to successfully meet the demands of the identified tasks?			
7. Can the level of abilities identified as essential for task performance be quantified?			
8. How does the System design affect skill requirements?			

	YES	NO	IN PROCESS
9. Has an optimal mix of skills among personnel been determined?			
10. Do ability/skill requirements change with operating and/or environmental conditions?			
11. How do reduced ability/skill levels impact upon System effectiveness?			

HOW MUCH MANPOWER IS REQUIRED?

	YES	NO	IN PROCESS
1. Have the tasks been structured into meaningful job units?			
2. Has an analysis of workload been conducted for each defined job?			
3. What is the optimal operator workload in terms of an increase in task performance and/or System effectiveness?			
4. Have functional relationships been established between System performance variables and required manpower?			
5. Have a number of different forecasting procedures been evaluated to			

	YES	NO	IN PROCESS
determine the best possible approach to quantifying manpower requirements?			
6. Is it possible to simulate the System based upon the task analysis in order to estimate manpower requirements?			
7. If the System cannot be simulated, are there statistical projections, expert opinions, or historical comparison data from which manpower requirements can be specified?			
8. How accurate are the manpower projections?			
9. Does the projected numbers of personnel possess the breadth of required skills?			

	YES	NO	IN PROCESS
10. Have number-skill trade-offs been considered?			
11. Have the manpower requirements been projected over the entire life-cycle of the System?			
12. Has a cost-effectiveness analysis been computed to determine total expense of the human resource component over the System's life-cycle?			
13. What is the impact of System design on the number of required personnel?			

WILL ADEQUATE NUMBERS OF APPROPRIATELY SKILLED PERSONNEL BE AVAILABLE?

	YES	NO	IN PROCESS
1. Have the specific sources (Army, Government agencies) providing data on manpower availability been determined?			
2. Have the MOS and EPMS been queried for availability of presently skilled personnel?			
3. What is the projection period of the manpower forecasts and does it span the life cycle of the System?			
4. Which forecasting method is best to use?			

	YES	NO	IN PROCESS
5. How reliable is a projection for this time period using the present forecasting method?			
6. Does the manpower availability forecast allow the level of accuracy needed to make valid decisions concerning the weapon System?			
7. Has the manpower availability forecast appropriately disaggregated the labor force into skills relevant to the System?			
8. Are the required skills presently available to the Army in sufficient numbers?			
9. If the required skills are not available, are they projected to be available when the System becomes operational?			

HOW CAN MANPOWER REQUIREMENTS BE MET?

	YES	NO	IN PROCESS
1. Is there an abundance or shortage of appropriately skilled manpower?			
2. Can manpower shortages be remedied by hardware--skill level--crew size--function trade-offs?			
3. Can manpower shortages be remedied by skill substitutability?			
4. Can manpower shortages be remedied by training?			
5. Can selection through recruitment remedy a shortage/overabundance problem?			
6. What is the impact of the remedies for shortages of manpower upon System effectiveness?			

	YES	NO	IN PROCESS
7. Can manpower shortages/over-abundance be avoided by incorporating human resource parameters early in System design?			

VI. DETAILED EXPLANATION OF CHAPTER 8

APPLIES TO: 1. Determining Which HPF's Should Be Diagnosed.

MARKING INADEQUATE NODES: If only performance values (P) have been computed, all nodes less than 50 are inadequate. If both performance values and upper confidence limit values (U) have been computed, all nodes with both values below 50 are inadequate. It is suggested that all such nodes be colored "red" as a visual aid.

DETERMINING THE SIGNIFICANCE OF NODES: Examine the color-coded nodes starting at the highest level of the Evaluation Tree, and determine if the performance at that node is sufficiently significant to require diagnosis. Significance of performance is determined by the impact that it has, or will have, on system evaluation. The higher in the Tree a node, the greater its impact on system evaluation. The following are <u>rules-of thumb</u> to be used as an aid in determining which nodes to diagnose:

(1) A System Node that is red should be diagnosed;
(2) A System Function Node that is red should be diagnosed;
(3) An SPI Node that is red should be diagnosed; and
(4) Any red node that is below the SPI level (and is not directly connected to a red SPI Node) does not have to be diagnosed. However, potentially useful information may result from diagnosing such nodes.

TRACING THROUGH THE EVALUATION TREE: To trace your way down the Evaluation Tree to reach the appropriate nodes for diagnosis, follow this procedure:

(1) Examine the Tree, and notice that all nodes are linked together by branches. Select the highest level red node on the Tree (that you want to diagnose).

(2) Follow the branches from this node to the node(s) directly below it that are red. There may be more than one such node.

(3) Repeat (1) and (2) until you reach the HPF Node level. The red nodes selected at this level are to be diagnosed.

APPLIES TO: 2. Determining which Diagnosis Measures to Apply to HPF's.

DIAGNOSTIC MEASURES: The two general types of diagnostic measures which apply to the constituent elements of the human resource areas are:

(1) Expert Measures. These are measures taken with the aid of one or more experts in the applicable human resource area. These measures consist of:

(a) relatively objective method for measuring the specific element, to be performed by the expert(s);
(b) method for structuring the opinions of the expert(s) about the relationship between the human resource area element and the selected HPF;
(c) method for structuring a combination of objective measurement and expert opinion about the element.

This type of measure provides data which will probably give you the best chance of producing a reasonably accurate diagnosis. However, this type of measure requires a significant amount of time and effort, and also may require resources which neither you nor your expert(s) will be able to obtain. This expert measure should be considered for use when time and resources permit, and when the circumstances demand the more accurate and thorough diagnosis available.

(2) Questionnaire Measures. These are measures based entirely on the opinions of the OT players and OT observers, as they apply to the selected HPF. This type of measure is based on the responses to the scales in the "Performance Difficulty Questionnaires" taken during the OT. This data has been summarized for increased usability in the scales of the "Opinion Summary Data Worksheet" for the HPF in question. There are two reasons for taking a questionnaire measure rather than an expert measure:

(a) Sufficient time and resources (including required data) are not available for taking the expert measure.

(b) After considering the parallel expert and questionnaire measures in detail, you prefer the methodology of the questionnaire measure, or you suspect that the opinions will give you more useful data.

There are three reasons for not taking a questionnaire measure:

(a) Because the participants and observers did not think that the given HPF was difficult to perform, the "Performance Difficulty Questionnaires" for this HPF were not filled out.

(b) Becaues the participants and observers did not think that the human resource elements measured by the Questionnaire scales were a source of problems in the performance of the given HPF, no mean scale scores are below 50.

(c) Since the questionnaire measures are based entirely on the subjective opinions of OT participants and observers, neither objective data nor expert opinions will be used to make the diagnosis for this specific human resource area element.